

Multi-class oscillating systems of interacting neurons

Susanne Ditlevsen *

Eva Löcherbach[†]

October 4, 2016

Abstract

We consider multi-class systems of interacting nonlinear Hawkes processes modeling several large families of neurons and study their mean field limits. As the total number of neurons goes to infinity we prove that the evolution within each class can be described by a nonlinear limit differential equation driven by a Poisson random measure, and state associated central limit theorems. We study situations in which the limit system exhibits oscillatory behavior, and relate the results to certain piecewise deterministic Markov processes and their diffusion approximations.

Key words : Multivariate nonlinear Hawkes processes. Mean-field approximations. Piecewise deterministic Markov processes. Multi-class systems. Oscillations. Diffusion approximation.

AMS Classification : 60G55; 60K35

1 Introduction

Biological rhythms are ubiquitous in living organisms. The brain controls and helps maintain the internal clock for many of these rhythms, and fundamental questions are how they arise and what is their purpose. Many examples of such biological oscillators can be found in the classical book by Glass and Mackey (1988) [17]. The motivation for this paper comes from the rhythmic scratch like network activity in the turtle, induced by a mechanical stimulus, and recorded and analyzed by Berg and co-workers [3, 4, 5, 27]. Oscillations in a spinal motoneuron are initiated by the sensory input, and continue by some internal mechanisms for some time after the stimulus is terminated. While mechanisms of rapid processing are well documented in sensory systems, rhythm-generating motor circuits in the spinal cord are poorly understood. The activation leads to an intense synaptic bombardment of both excitatory and inhibitory input, and it is of interest to characterize such network activity, and to build models which can generate self-sustained oscillations.

The aim of this paper is to present a microscopic model describing a large network of interacting neurons which can generate oscillations. The activity of each neuron is represented by a point process, namely, the successive times at which the neuron emits an

*Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen

[†]Université de Cergy-Pontoise, AGM UMR-CNRS 8088, 2 avenue Adolphe Chauvin, F-95302 Cergy-Pontoise Cedex

action potential or a so-called spike. A realization of this point process is called a spike train. It is commonly admitted that the spiking intensity of a neuron, i.e., the infinitesimal probability of emitting an action potential during the next time unit, depends on the past history of the neuron and it is affected by the activity of other neurons in the network. Neurons interact mostly through chemical synapses, where a spike of a pre-synaptic neuron leads to an increase if the synapse is excitatory, or a decrease if the synapse is inhibitory, of the membrane potential of the post-synaptic neuron, possibly after some delay. In neurophysiological terms this is called synaptic integration. When the membrane potential reaches a certain upper threshold, the neuron fires a spike. Thus, excitatory inputs from the neurons in the network increase the firing intensity, and inhibitory inputs decrease it. Hawkes processes provide good models of this synaptic integration phenomenon by the structure of their intensity processes, see (1.1) below. We refer to Chevallier et al. (2015) [8], Chornoboy et al. (1988) [9], Hansen et al. (2015) [20] and to Reynaud-Bouret et al. (2014) [34] for the use of Hawkes processes in neuronal modeling. For an overview of point processes used as stochastic models for interacting neurons both in discrete and in continuous time and related issues, see also Galves and Löcherbach (2016) [16].

In this paper, we study oscillatory systems of interacting Hawkes processes representing the time occurrences of action potentials of neurons. The system consists of several large populations of neurons. Each population might represent a different functional group of neurons, for example different hierarchical layers in the visual cortex, such as V1 to V4, or the populations can be pools of excitatory and inhibitory neurons in a network. Each neuron is characterized by its spike train, and the whole system is described by multivariate counting processes $Z_{k,i}^N(t), t \geq 0$. Here, $Z_{k,i}^N(t)$ represents the number of spikes of the i th neuron belonging to the k th population, during the time interval $[0, t]$. The number of classes n is fixed, and each class $k = 1, \dots, n$ consists of N_k neurons. The total number of neurons is therefore $N = N_1 + \dots + N_n$.

Under suitable assumptions, the sequence of counting processes $(Z_{k,i}^N)_{1 \leq k \leq n, 1 \leq i \leq N_k}$ is characterized by its intensity processes $(\lambda_{k,i}^N(t))$ defined through the relation

$$\mathbb{P}(Z_{k,i}^N \text{ has a jump in }]t, t + dt| \mathcal{F}_t) = \lambda_{k,i}^N(t) dt,$$

where $\mathcal{F}_t = \sigma(Z_{k,i}^N(s), s \leq t, 1 \leq k \leq n, 1 \leq i \leq N_k)$. We consider a mean-field framework where $\lambda_{k,i}^N(t)$ is given by

$$\lambda_{k,i}^N(t) = f_k \left(\sum_{l=1}^n \frac{1}{N_l} \sum_{1 \leq j \leq N_l} \int_{]0, t[} h_{kl}(t-s) dZ_{l,j}^N(s) \right). \quad (1.1)$$

Here, $f_k : \mathbb{R} \rightarrow \mathbb{R}_+$ is the *spiking rate function* of population k , and $\{h_{kl} : \mathbb{R}_+ \rightarrow \mathbb{R}\}$ is a family of *synaptic weight functions* modeling the influence of population l on population k . By integrating over $]0, t[$ and not over $] - \infty, t[$, we implicitly assume initial conditions of no spiking activity before time 0.

Equation (1.1) has the typical form of the intensity of a multivariate nonlinear Hawkes process, going back to Hawkes (1971) [21] and Hawkes and Oakes (1974) [22]. We refer to Brémaud and Massoulié (1996) [6] for the stability properties of multivariate Hawkes processes, and to Delattre, Fournier and Hoffmann (2015) [13] and Chevallier (2015) [7] for the study of Hawkes processes in high dimensions.

The structure of (1.1) is such that within each population, all neurons behave in a similar way, i.e., the intensity process $\lambda_{k,i}^N(t)$ depends only on the empirical measures of each population. Thus, neurons within a given population are exchangeable. Therefore, we deal with a multi-class system of populations interacting in a mean-field framework which is reminiscent of Graham (2008) [18] and Graham and Robert (2009) [19]. Our aim is to study the large population limit when $N \rightarrow \infty$ and to show that in this limit self-sustained periodic behavior emerges even though each single neuron does not follow periodic dynamics. The study follows a long tradition, see e.g. Scheutzow (1985) [31, 32] in the framework of nonlinear diffusion processes, or Dai Pra, Fischer and Regoli (2015) [12] and Collet, Dai Pra and Formentin (2015) [10]. Our paper continues these studies within the framework of infinite memory point processes. The first important step is to establish propagation of chaos of the finite system $(Z_{k,i}^N(t))_{1 \leq k \leq n, 1 \leq i \leq N_k}$ as $N \rightarrow \infty$, under the condition that for each class $1 \leq k \leq n$, $\lim_{N \rightarrow \infty} N_k/N$ exists and is in $]0, 1[$.

1.1 Propagation of chaos

In Section 2.2, we study the limit behavior of the system $(Z_{k,i}^N(t))_{1 \leq k \leq n, 1 \leq i \leq N_k}$ as $N \rightarrow \infty$. We show in Theorem 1 that the system can be approximated by a system of inhomogeneous independent Poisson processes $(\bar{Z}_1(t), \dots, \bar{Z}_n(t))$, where each $\bar{Z}_k(t)$ has intensity

$$f_k \left(\sum_{l=1}^n \int_0^t h_{kl}(t-s) d\mathbb{E}(\bar{Z}_l(s)) \right) dt.$$

Here, $\bar{Z}_k(t)$ represents the number of spikes during $[0, t]$ of a typical neuron belonging to population k in the limit system. This result is an extension of results obtained by [13] to the multi-class case. It follows that the system is *multi-chaotic* in the sense of [18]. The equivalence between the *chaoticity* of the system and a weak law of large numbers for the empirical measures, as proven in Theorem 1, is well-known (see for instance Sznitman (1991) [36]). This means that in the large population limit, within the same class, the neurons converge in law to independent and identically distributed copies of the same limit law. This property is usually called *propagation of chaos* in the literature. In particular, as pointed out in [18], we have asymptotic independence between the different classes, and interactions between classes do only survive in law.

In Section 3, still following ideas of [13], we state an associated central limit theorem in Theorem 2. The extension to nonlinear rate functions f_k requires the use of matrix-convolution equations which go back to Crump (1970) [11] and Athreya and Murthy (1976) [1] and which are collected in Appendix, Section 6.1.

1.2 Oscillatory behavior of the limit system

In Section 4 we present conditions under which the limit system possesses solutions which are periodic in law. To be more precise, the classes interact according to a cyclic feedback system and each class k is only influenced by class $k+1$, where we identify $n+1$ with 1. In this case $m_t^k = \mathbb{E}(\bar{Z}_k(t))$, $1 \leq k \leq n$, is solution of

$$m_t^k = \int_0^t f_k \left(\int_0^s h_{kk+1}(s-u) dm_u^{k+1} \right) ds. \quad (1.2)$$

If the memory kernels h_{kk+1} are given by Erlang kernels, as used e.g. in modeling the delay in the hemodynamics in nephrons, see [14, 35] and (4.17) below, then Theorem 3 characterizes situations in which the system (1.2) possesses attracting non-constant periodic orbits, that is, presents oscillatory behavior. This result goes back to deep theorems in dynamical systems, obtained by Mallet-Paret and Smith (1990) [30] and used in a different context in Benaïm and Hirsch (1999) [2], from where we learned about these results. In particular, the celebrated Poincaré-Bendixson theorem plays a crucial role.

1.3 Hawkes processes, associated piecewise deterministic Markov processes and longtime behavior of the approximating diffusion process

Hawkes processes are truly infinite memory processes and techniques from the theory of Markov processes are in general not applicable. However, in the special situation where the memory kernels are given by Erlang kernels, the intensity processes can be described in terms of an equivalent high dimensional system of *piecewise deterministic Markov processes* (PDMPs). Once we are back in the Markovian world, we can study the longtime behavior of the process, ergodicity, and so on.

In Section 5, we obtain an approximating diffusion equation in (5.26) which is shown to be close in a weak sense to the original PDMP defining the Hawkes process (Theorem 4). Once we dispose of this small noise approximation, we then study the longtime behavior in the case of two populations, $n = 2$. In particular, we show to which extent the approximating diffusion presents the same oscillatory behavior as the limit system.

This approximating diffusion is highly degenerate having Brownian noise present only in two of its coordinates. However, the very specific *cascade structure* of its drift vector implies that the weak Hörmander condition holds on the whole state space, and as a consequence, the diffusion is strong Feller. A simple Lyapunov argument shows that the process comes back to a compact set infinitely often, almost surely.

Since the limit system possesses a non constant periodic orbit Γ which is asymptotically orbitally stable, it is well known that there exists a local Lyapunov function $V(x)$ defined on a neighborhood of Γ such that V decreases along the trajectories of the limit system, describing the attraction of the limit system to Γ (see e.g. Yoshizawa (1966) [38] and Kloeden and Lorenz (1986) [28]). This Lyapunov function is shown also to be a Lyapunov function for the approximating diffusion, in particular, the diffusion is also attracted to Γ , once it has entered the basin of attraction of Γ . A control argument shows finally that this happens infinitely often almost surely (Theorem 5), in particular, for large enough N , the approximating diffusion also presents oscillations.

We close our paper with some simulation studies.

2 Systems of interacting Hawkes processes, basic notation and large population limits

Consider n populations, each composed by N_k neurons, $k = 1, \dots, n$. The total number of neurons in the system is $N = N_1 + \dots + N_n$. The activity of each neuron is described by a counting process $Z_{k,i}^N(t)$, $1 \leq k \leq n$, $1 \leq i \leq N_k$, $t \geq 0$, recording the number of spikes of the i th neuron belonging to population k during the interval $[0, t]$. The sequence of

counting processes $(Z_{k,i}^N)$ is characterized by its intensity processes $(\lambda_{k,i}^N(t))$ which are defined through the relation

$$\mathbb{P}(Z_{k,i}^N \text{ has a jump in }]t, t+dt] | \mathcal{F}_t) = \lambda_{k,i}^N(t) dt, 1 \leq k \leq n, 1 \leq i \leq N_k,$$

where $\mathcal{F}_t = \sigma(Z_{k,i}^N(s), s \leq t, 1 \leq k \leq n, 1 \leq i \leq N_k)$ and $\lambda_{k,i}^N(t)$ are defined in (2.3) below. We consider a mean field framework where $N \rightarrow \infty$ such that for each $1 \leq k \leq n$,

$$\lim_{N \rightarrow \infty} \frac{N_k}{N} = p_k \text{ exists and is in }]0, 1[.$$

The intensity processes will be of the form

$$\lambda_{k,i}^N(t) = f_k \left(\sum_{l=1}^n \frac{1}{N_l} \sum_{1 \leq j \leq N_l} \int_{]0, t[} h_{kl}(t-s) dZ_{l,j}^N(s) \right), \quad (2.3)$$

where f_k is the spiking rate function of population k and where the h_{kl} are memory kernels.

Assumption 1 (i) All f_k belong to $C^1(\mathbb{R}; \mathbb{R}_+)$.

(ii) There exists a finite constant L such that for every x and x' in \mathbb{R} , for every $1 \leq k \leq n$,

$$|f_k(x) - f_k(x')| \leq L|x - x'|. \quad (2.4)$$

(iii) The functions $h_{kl}, 1 \leq k, l \leq n$, belong to $L_{loc}^2(\mathbb{R}_+; \mathbb{R})$.

2.1 The setting

We work on a filtered probability space $(\Omega, \mathcal{A}, \mathbb{F})$ which we define as follows. We write \mathbb{M} for the canonical path space of simple point processes given by

$$\mathbb{M} := \{\mathbf{m} = (t_n)_{n \in \mathbb{N}} : t_1 > 0, t_n \leq t_{n+1}, t_n < t_{n+1} \text{ if } t_n < +\infty, \lim_{n \rightarrow +\infty} t_n = +\infty\}.$$

For any $\mathbf{m} \in \mathbb{M}$, any $n \in \mathbb{N}$, let $T_n(\mathbf{m}) = t_n$. We identify $\mathbf{m} \in \mathbb{M}$ with the associated point measure $\mu = \sum_n \delta_{T_n(\mathbf{m})}$ and put $\mathcal{M}_t := \sigma\{\mu(A) : A \in \mathcal{B}(\mathbb{R}), A \subset [0, t]\}$, $\mathcal{M} = \mathcal{M}_\infty$. Finally, we put $(\Omega, \mathcal{A}, \mathbb{F}) := (\mathbb{M}, \mathcal{M}, (\mathcal{M}_t)_{t \geq 0})^I$ where $I = \bigcup_{k=1}^n \{(k, i), i \geq 1\}$. We write $(Z_{k,i}^N)_{1 \leq k \leq n, 1 \leq i \leq N_k}$ for the canonical multivariate point measure defined on the finite dimensional subspace $(\mathbb{M}, \mathcal{M}, (\mathcal{M}_t)_{t \geq 0})^{I^N}$ of Ω , where $I^N = \bigcup_{k=1}^n \{(k, i), 1 \leq i \leq N_k\}$.

Definition 1 [compare to Definition 1 of [13]].

A Hawkes process with parameters $(f_k, h_{kl}, 1 \leq k, l \leq n)$ is a probability measure \mathbb{P} on $(\Omega, \mathcal{A}, \mathbb{F})$ such that

1. \mathbb{P} -almost surely, for all $(k, i) \neq (l, j)$, $Z_{k,i}^N$ and $Z_{l,j}^N$ never jump simultaneously,
2. for all $(k, i) \in I^N$, the compensator of $Z_{k,i}^N(t)$ is given by $\int_0^t \lambda_{k,i}^N(s) ds$ defined in (2.3).

Proposition 1 Under Assumption 1 there exists a path-wise unique Hawkes process $(Z_{k,i}^N(t))_{(k,i) \in I^N}$ for all $t \geq 0$.

Proof The proof is analogous to the proof of Theorem 6 in [13]. •

2.2 Mean-field limit and propagation of chaos

The aim of the paper is to study the process $(Z_{k,i}^N(t))_{(k,i) \in I^N}$ in the large population limit, i.e., as $N \rightarrow \infty$. The convergence will be stated in terms of the empirical measures

$$\frac{1}{N_k} \sum_{1 \leq i \leq N_k} \delta_{(Z_{k,i}^N(t))_{t \geq 0}}, \quad 1 \leq k \leq n, \quad (2.5)$$

taking values in the set $\mathcal{P}(D(\mathbb{R}_+, \mathbb{R}_+))$ of probability measures on the space of càdlàg functions, $D(\mathbb{R}_+, \mathbb{R}_+)$. We endow $D(\mathbb{R}_+, \mathbb{R}_+)$ with the Skorokhod topology, and $\mathcal{P}(D(\mathbb{R}_+, \mathbb{R}_+))$ with the weak convergence topology associated with the Skorokhod topology on $D(\mathbb{R}_+, \mathbb{R}_+)$.

Since we are dealing with multi-class systems, the classical notions of chaoticity and propagation of chaos have to be extended to this framework, see [18] for further details. We recall from [18] the following definition. Let $P_1, \dots, P_n \in \mathcal{P}(D(\mathbb{R}_+, \mathbb{R}_+))$.

Definition 2 *The system $(Z_{k,i}^N(t))_{(k,i) \in I^N}$ is called $P_1 \otimes \dots \otimes P_n$ -multi-chaotic, if for any $m \geq 1$,*

$$\lim_{N \rightarrow \infty} \mathcal{L}((Z_{k,i}^N(t)), 1 \leq k \leq n, 1 \leq i \leq m) = P_1^{\otimes m} \otimes \dots \otimes P_n^{\otimes m}.$$

In particular, Corollary 5.2 of [19] shows that in this case we have convergence in distribution

$$\frac{1}{N_k} \sum_{1 \leq i \leq N_k} \delta_{(Z_{k,i}^N(t))_{t \geq 0}} \xrightarrow{\mathcal{L}} P_k,$$

as $N \rightarrow \infty$, for any $1 \leq k \leq n$. The limit measure P_k has to be understood as the distribution of the limit process $\bar{Z}_k(t)$, where the associated limit system is given by

$$\bar{Z}_k(t) = \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{z \leq f_k(\sum_{l=1}^n \int_0^s h_{kl}(s-u) d\mathbb{E}(\bar{Z}_l(u))\}} N^k(ds, dz), \quad (2.6)$$

$1 \leq k \leq n$, where N^k are independent Poisson random measures (PRMs) on $\mathbb{R}_+ \times \mathbb{R}_+$ each having intensity measure $dsdz$.

Introduce $m_t = (m_t^1, \dots, m_t^n) = (\mathbb{E}(\bar{Z}_1(t)), \dots, \mathbb{E}(\bar{Z}_n(t)))$. Taking expectations in (2.6), it follows that m_t is solution of

$$m_t^k = \int_0^t f_k \left(\sum_{l=1}^n \int_0^s h_{kl}(s-u) dm_u^l \right) ds, \quad 1 \leq k \leq n. \quad (2.7)$$

Theorem 1 *Under Assumption 1, there exists a path-wise unique solution to (2.6) such that $t \mapsto \mathbb{E}(\sum_{k=1}^n \bar{Z}_k(t))$ is locally bounded. Moreover, the system of processes $(Z_{k,i}^N(t))_{(k,i) \in I^N}$ is $P_1 \otimes \dots \otimes P_n$ -multi-chaotic, where $P_k = \mathcal{L}(\bar{Z}_k)$, $1 \leq k \leq n$. In particular, for any $i \geq 1$,*

$$((Z_{1,i}^N(t), \dots, Z_{n,i}^N(t))_{t \geq 0}) \xrightarrow{\mathcal{L}} ((\bar{Z}_1(t), \dots, \bar{Z}_n(t))_{t \geq 0})$$

as $N \rightarrow \infty$ (convergence in $D(\mathbb{R}_+, \mathbb{R}_+^n)$, endowed with the Skorokhod topology).

Remark 1 *The above theorem shows that any fixed finite sub-system is asymptotically independent with neurons of class k having the law of \bar{Z}_k .*

The proof of Theorem 1 is a direct adaptation of the proof of Theorem 8 in [13] to the multi-class case.

Proof 1) Let $(\bar{Z}_1(t), \dots, \bar{Z}_n(t))$ be any solution of (2.6) and consider the associated vector $m_t = (m_t^1, \dots, m_t^n) = (\mathbb{E}(\bar{Z}_1(t), \dots, \bar{Z}_n(t)))$. Then m_t is solution of (2.7), and an easy adaptation of Lemma 24 of [13] shows that this equation has a unique non-decreasing (in each coordinate) locally bounded solution, which is of class C^1 .

2) Well-posedness and uniqueness of a solution satisfying that $t \mapsto \mathbb{E}(\sum_{k=1}^n \bar{Z}_k(t))$ is locally bounded follow then as in [13], proof of Theorem 8.

3) Propagation of chaos: Let $N_{k,i}^N(ds, dz), (k, i) \in I^N$, be i.i.d. PRMs having intensity $dsdz$ on $\mathbb{R}_+ \times \mathbb{R}_+$. For each $N \geq 1$, consider the Hawkes process $(Z_{k,i}^N(t))_{(k,i) \in I^N, t \geq 0}$ given by

$$\begin{aligned} Z_{k,i}^N(t) &= \int_0^t \int_0^\infty \mathbf{1}_{\{z \leq \lambda_{k,i}^N(s)\}} N_{k,i}^N(ds, dz) \\ &= \int_0^t \int_0^\infty \mathbf{1}_{\{z \leq f_k(\sum_{l=1}^n \frac{1}{N_l} \sum_{1 \leq j \leq N_l} \int_0^{s-} h_{kl}(s-u) dZ_{l,j}^N(u))\}} N_{k,i}^N(ds, dz). \end{aligned}$$

Indeed, $Z_{k,i}^N$ defined in this way is a Hawkes process in the sense of Definition 1, as follows from Proposition 3 of [13]. We now couple $Z_{k,i}^N$ with the limit process (2.6) in the following way. Let m_t be the unique solution of (2.7). Put

$$\bar{Z}_{k,i}^N(t) = \int_0^t \int_0^\infty \mathbf{1}_{\{z \leq f_k(\sum_{l=1}^n \int_0^s h_{kl}(s-u) dm_u^l)\}} N_{k,i}^N(ds, dz), \quad (2.8)$$

where $N_{k,i}^N$ is the PRM driving the dynamics of $Z_{k,i}^N$. Obviously, for all $1 \leq i \leq N_k$, $\bar{Z}_{k,i}^N \stackrel{\mathcal{L}}{=} \bar{Z}_k$. Moreover, the limit processes $\bar{Z}_{k,i}^N, 1 \leq k \leq n, 1 \leq i \leq N_k$, are independent.

Denote $\Delta_{k,i}^N(t) = \int_0^t |d[\bar{Z}_{k,i}^N(u) - Z_{k,i}^N(u)]|$, and $\delta_{k,i}^N(t) = \mathbb{E}(\Delta_{k,i}^N(t))$. Notice that this last quantity does not depend on $i \in \{1, \dots, N_k\}$, due to the exchangeability of the neurons within one class. Then

$$\sup_{u \in [0, t]} |\bar{Z}_{k,i}^N(u) - Z_{k,i}^N(u)| \leq \Delta_{k,i}^N(t), \text{ whence } \mathbb{E} \left[\sup_{u \in [0, t]} |\bar{Z}_{k,i}^N(u) - Z_{k,i}^N(u)| \right] \leq \delta_{k,1}^N(t) := \delta_k^N(t).$$

We start by controlling $\Delta_{k,i}^N(t)$ which is given by

$$\begin{aligned} \Delta_{k,i}^N(t) &= \int_0^t \int_0^\infty \left| \mathbf{1}_{\{z \leq f_k(\sum_{l=1}^n \int_0^s h_{kl}(s-u) dm_u^l)\}} \right. \\ &\quad \left. - \mathbf{1}_{\{z \leq f_k(\sum_{l=1}^n \frac{1}{N_l} \sum_{1 \leq j \leq N_l} \int_0^{s-} h_{kl}(s-u) dZ_{l,j}^N(u))\}} \right| N_{k,i}^N(ds, dz). \end{aligned}$$

Using the Lipschitz continuity of f_k with Lipschitz constant L ,

$$\begin{aligned} \frac{1}{L} \mathbb{E}(\Delta_{k,i}^N(t)) &\leq \int_0^t \sum_{l=1}^n \mathbb{E} \left| \frac{1}{N_l} \sum_{1 \leq j \leq N_l} \int_0^{s-} h_{kl}(s-u) (dm_u^l - d\bar{Z}_{l,j}^N(u)) \right| ds \\ &\quad + \int_0^t \sum_{l=1}^n \mathbb{E} \left| \frac{1}{N_l} \sum_{1 \leq j \leq N_l} \int_0^{s-} h_{kl}(s-u) d[\bar{Z}_{l,j}^N(u) - Z_{l,j}^N(u)] \right| ds =: A + B, \quad (2.9) \end{aligned}$$

where A denotes the terms of the RHS of the first line, and B the terms within the second line. Now, using Lemma 22 of [13],

$$B \leq \int_0^t \mathbb{E} \int_0^{s-} \left[\sum_{l=1}^n |h_{kl}(s-u)| d\Delta_{l,1}^N(u) \right] ds \leq \int_0^t \left[\sum_{l=1}^n |h_{kl}(t-u)| \delta_l^N(u) \right] du.$$

To control A , let $X_{k,l,j}^N(t) = \int_0^{t-} h_{kl}(t-u) d\bar{Z}_{l,j}^N(u)$, for $1 \leq j \leq N_l$. Then $X_{k,l,j}^N(t)$, $1 \leq j \leq N_l$, are i.i.d. having mean $\int_0^t h_{kl}(t-u) dm_u^l$. Hence

$$A \leq \sum_{l=1}^n \frac{1}{\sqrt{N_l}} \int_0^t \sqrt{\text{Var}(X_{k,l,1}^N(s))} ds.$$

But

$$X_{k,l,1}^N(s) = \int_0^{s-} \int_0^\infty \mathbf{1}_{\{z \leq f_l(\sum_{m=1}^n \int_0^u h_{lm}(u-r) dm_r^m)\}} h_{kl}(s-u) N_{l,1}^N(du, dz),$$

and thus, since the integrand is deterministic,

$$X_{k,l,1}^N(s) - \mathbb{E}(X_{k,l,1}^N(s)) = \int_0^{s-} \int_0^\infty \mathbf{1}_{\{z \leq f_l(\sum_{m=1}^n \int_0^u h_{lm}(u-r) dm_r^m)\}} h_{kl}(s-u) \tilde{N}_{l,1}^N(du, dz),$$

where $\tilde{N}_{l,1}^N(ds, dz) = N_{l,1}^N(ds, dz) - ds dz$ is the compensated PRM. Recalling (2.7) we deduce that

$$\text{Var}(X_{k,l,1}^N(s)) = \int_0^s f_l \left(\sum_{m=1}^n \int_0^u h_{lm}(u-r) dm_r^m \right) h_{kl}^2(s-u) du = \int_0^s h_{kl}^2(s-u) dm_u^l.$$

Putting $\|\delta^N(t)\|_1 = \sum_{k=1}^n \delta_k^N(t)$, $\|m_t\|_1 = \sum_{k=1}^n m_t^k$, $\|h(t)\|_1 = \sum_{k,l=1}^n |h_{kl}(t)|$, we obtain

$$\begin{aligned} \frac{1}{L} \|\delta^N(t)\|_1 &\leq \int_0^t \|h(t-u)\|_1 \|\delta^N(u)\|_1 du \\ &\quad + \left(\sum_{k=1}^n \frac{1}{\sqrt{N_k}} \right) \int_0^t \left(\int_0^s \|h(s-u)\|_1^2 dm_u \right)^{1/2} ds. \end{aligned} \quad (2.10)$$

It follows, as in Step 4 of the proof of Theorem 8 in [13], that

$$\sup_{t \leq T} \|\delta^N(t)\|_1 \leq C_T \left(\sum_{k=1}^n \frac{1}{\sqrt{N_k}} \right).$$

Consequently, for any fixed $(k, i) \in I^N$, as $N \rightarrow \infty$,

$$\mathbb{E} \left[\sup_{u \in [0, T]} |\bar{Z}_{k,i}^N(u) - Z_{k,i}^N(u)| \right] \leq C_T \left(\sum_{k=1}^n \frac{1}{\sqrt{N_k}} \right) \rightarrow 0. \quad (2.11)$$

The end of the proof is now standard, based on arguments developed in [18] and [19]. Recall that neurons within a given population are exchangeable. Then, by the proof of Theorem 5.1 in [19], in order to prove propagation of chaos, it is enough to show that for each fixed sequence ℓ_k , $1 \leq k \leq n$,

$$((Z_{1,1}^N(t))_{t \geq 0}, \dots, (Z_{1,\ell_1}^N(t))_{t \geq 0}, \dots, (Z_{n,1}^N(t))_{t \geq 0}, \dots, (Z_{n,\ell_n}^N(t))_{t \geq 0})$$

goes in law to ℓ_1 independent copies of \bar{Z}_1, \dots , and ℓ_n independent copies of \bar{Z}_n (convergence in $D(\mathbb{R}_+, \mathbb{R}_+^{\ell_1 + \dots + \ell_n})$). Since the topology of uniform convergence on compact time intervals is finer than the Skorokhod topology, this follows clearly from (2.11), and thus the proof is finished. \bullet

3 Central limit theorem

A natural question to ask is to which extent the large time behavior of the limit system (m_t^1, \dots, m_t^n) predicts the large time behavior of the finite size system, in particular in the case when the limit system presents oscillations (see Section 4 below). To answer this question, the present section states a central limit theorem where convergence of both N and t to infinity is considered. All proofs can be found in Appendix.

First we control the longtime behavior of the limit system represented by its integrated intensities (m_t^1, \dots, m_t^n) . It is well-known that linear Hawkes processes, i.e., the case when the rate functions f_k are linear, can be described in terms of classical Galton-Watson processes. In the non-linear case, a comparison with a Galton-Watson process is still possible if the rate functions are Lipschitz. In our case, the associated offspring matrix is given by $\Lambda := (\Lambda_{ij})_{1 \leq i, j \leq n}$, where

$$\Lambda_{ij} = L \int_0^\infty |h_{ij}(t)| dt, \quad 1 \leq i, j \leq n, \quad (3.12)$$

with L given in (2.4). Define the matrix

$$H(t) = \left(L |h_{ik}(t)| \right)_{1 \leq i, k \leq n}, \quad \text{for any } t \geq 0, \quad (3.13)$$

such that $\Lambda = \int_0^\infty H(t) dt$.

Classically, one distinguishes the subcritical, the critical and the supercritical cases. Since we only need to bound the intensities, we concentrate on the subcritical and the supercritical case. The subcritical case is defined by the following property of the matrix Λ .

Assumption 2 *The functions $h_{kl}, 1 \leq k, l \leq n$, belong to $L^1(\mathbb{R}_+; \mathbb{R}) \cap L^2(\mathbb{R}_+; \mathbb{R})$, and the largest eigenvalue μ_1 of Λ is strictly smaller than 1.*

We then obtain the following bound on the growth of m_t^k .

Proposition 2 *Grant Assumptions 1 and 2. Then there exists a constant α_0 such that*

$$m_t^k \leq \alpha_0 t, \quad \text{for all } 1 \leq k \leq n.$$

Moreover, as in [13], Remark 9, in this case there exists a constant C such that

$$\mathbb{E}(\sup_{s \leq t} |Z_{k,i}^N(s) - \bar{Z}_{k,i}^N(s)|) \leq CtN^{-1/2}, \quad (3.14)$$

for $1 \leq k \leq n$, where $\bar{Z}_{k,i}^N$ is defined in (2.8).

In the supercritical case, the control on the growth of m_t^k is more tricky. We need the following assumption.

Assumption 3 The functions $h_{kl}, 1 \leq k, l \leq n$, belong to $L^1(\mathbb{R}_+; \mathbb{R})$ and there exist $p \geq 1$ and a constant C , such that $|h_{kl}(t)| \leq C(1 + t^p)$ for all $t \geq 0$, for any $1 \leq k, l \leq n$. Moreover, the largest eigenvalue μ_1 of Λ in (3.12) is strictly larger than 1.

Proposition 3 Grant Assumptions 1 and 3. Then for all $1 \leq k \leq n$,

$$m_t^k \leq ce^{\alpha_0 t},$$

where α_0 is unique such that $\int_0^\infty e^{-\alpha_0 t} H(t) dt$ has largest eigenvalue $\equiv +1$. Here, $H(t)$ is given in (3.13). Moreover, for $1 \leq k \leq n$ and for some constant C ,

$$\mathbb{E}(\sup_{s \leq t} |Z_{k,i}^N(s) - \bar{Z}_{k,i}^N(s)|) \leq Ce^{\alpha_0 t} N^{-1/2}. \quad (3.15)$$

We obtain the following central limit theorem. It is an extension of Theorem 10 of [13] to the nonlinear case and several populations.

Theorem 2 Grant Assumption 1 and either Assumption 2 or 3. Suppose moreover that for all $1 \leq k \leq n$, $\liminf_{t \rightarrow \infty} m_t^k/t \geq \alpha_k$ for some $\alpha_k > 0$. We will consider limits as both N and t tend to infinity, under the constraints that $t/N \rightarrow 0$ in the subcritical case and that $e^{\alpha_0 t} t^{-1} N^{-1/2} \rightarrow 0$ in the supercritical case, where α_0 is given in Proposition 3.

1. For any fixed i , we have that $Z_{k,i}^N(t)/m_t^k$ tends to 1 in probability. More precisely,

$$\lim_{N, t \rightarrow \infty} \sup (m_t^k)^{1/2} \mathbb{E}[|Z_{k,i}^N(t)/m_t^k - 1|] \leq C,$$

for some constant C and for $k = 1, \dots, n$.

2. For any fixed ℓ_1, \dots, ℓ_n , the vector

$$\left(\left(\frac{Z_{1,i}^N(t) - m_t^1}{\sqrt{m_t^1}} \right)_{1 \leq i \leq \ell_1}, \dots, \left(\frac{Z_{n,i}^N(t) - m_t^n}{\sqrt{m_t^n}} \right)_{1 \leq i \leq \ell_n} \right)$$

tends in law to $\mathcal{N}(0, I_{\ell_1 + \dots + \ell_n})$ as $(t, N) \rightarrow \infty$, under the constraint that $\lim_{N \rightarrow \infty} N_k/N > 0$ for all $1 \leq k \leq n$.

The proof of Theorem 2 and Propositions 2 and 3 can be found in Appendix.

Remark 2 Since the rate functions are nonlinear, we only obtain the central limit theorem in the regime $tN^{-1} \rightarrow 0$ (subcritical case) or $N^{-1/2} t^{-1} e^{\alpha_0 t} \rightarrow 0$ (supercritical case), contrarily to [13] who do not have any restriction in the subcritical case and who only impose $N^{-1} e^{\alpha_0 t} \rightarrow 0$ in the supercritical case. This is due to the fact on the one hand that we deal with the nonlinear case and on the other hand that we do not dispose of general asymptotical equivalents of $t \mapsto m_t^k$.

4 Oscillations and associated dynamical systems in monotone cyclic feedback systems

The aim of this section is to study the limit system (2.6) and (2.7) and describe situations in which oscillations will occur. Throughout this section we suppose that the information is transported through the system according to a *monotone cyclic feedback system* [30]. That it is monotone means that the rate functions f_k are non-decreasing, and a cyclic feedback system means that each population k is only influenced by population $k + 1$, where we identify $n + 1$ with 1. Thus, $h_{kl} \equiv 0$ for all k, l such that $l \neq k + 1$. The memory kernels h_{kk+1} describe how population $k + 1$ influences population k .

From now on, we identify m^{n+1} with m^1 and introduce the memory variables

$$x_t^k = \int_0^t h_{kk+1}(t-s) dm_s^{k+1}, \text{ for } 1 \leq k \leq n. \quad (4.16)$$

We have $m_t^k = \int_0^t f_k(x_s^k) ds$. For specific choices of kernel functions the above system of memory variables can be developed into a system of differential equations without delay by increasing the dimension of the system, see (4.19) below. We call this a *Markovian cascade* of successive memory terms. It is obtained by using Erlang kernels, given by

$$h_{kk+1}(s) = c_k e^{-\nu_k s} \frac{s^{\eta_k}}{\eta_k!}, \text{ for } k < n, \text{ and } h_{n1}(s) = c_n e^{-\nu_n s} \frac{s^{\eta_n}}{\eta_n!}, \quad (4.17)$$

where $\eta_k \in \mathbb{N}_0, c_k \in \{-1, 1\}$ and $\nu_k > 0$ are fixed constants. Here, $\eta_k + 1$ is the order of the delay, i.e., the number of differential equations needed for population k to obtain a system without delay terms. The delay of the influence of population $k + 1$ on population k is distributed and taking its maximum absolute value at η_k/ν_k time units back in time, and the mean is $(\eta_k + 1)/\nu_k$ (if normalizing to a probability density). The higher the order of the delay, the more concentrated is the delay around its mean value, and in the limit of $\eta_k \rightarrow \infty$ while keeping $(\eta_k + 1)/\nu_k$ fixed, the delay converges to a discrete delay. The sign of c_k indicates if the influence is inhibitory or excitatory.

Observing that $h'_{kk+1}(t) = -\nu_k h_{kk+1}(t) + c_k \frac{t^{\eta_k-1}}{(\eta_k-1)!} e^{-\nu_k t}$ leads to the following auxiliary variables

$$x_t^{k,l} = \int_0^t c_k e^{-\nu_k(t-s)} \frac{(t-s)^{\eta_k-l}}{(\eta_k-l)!} dm_s^{k+1}, 1 \leq k \leq n, 0 \leq l \leq \eta_k,$$

where we identify $x^k = x^{k,0}$. Then we can rewrite

$$\frac{dx_t^{k,l}}{dt} = -\nu_k x_t^{k,l} + x_t^{k,l+1}, l < \eta_k. \quad (4.18)$$

Iterating this argument, the following system of coupled differential equations is obtained. For all $1 \leq k \leq n$, and where as usual $n + 1$ is identified with 1,

$$\begin{aligned} \frac{dx_t^{k,l}}{dt} &= -\nu_k x_t^{k,l} + x_t^{k,l+1}, 0 \leq l < \eta_k, \\ \frac{dx_t^{k,\eta_k}}{dt} &= -\nu_k x_t^{k,\eta_k} + c_k f_{k+1}(x_t^{k+1,0}), \end{aligned} \quad (4.19)$$

with initial conditions $x_0^{k,l} = 0$. System (4.19) exhibits the structure of a monotone cyclic feedback system as considered e.g. in [30] or as (33) and (34) in [2]. If $\prod_{k=1}^n c_k > 0$, then

the system (4.19) is of total positive feedback, otherwise it is of negative feedback. We obtain the following simple first result.

Proposition 4 *Suppose that $\prod_{k=1}^n c_k < 0$ and that f_1, \dots, f_n are non-decreasing. Then (4.19) admits a unique equilibrium x^* .*

Proof Any equilibrium x^* must satisfy

$$(x^*)^{n,\eta_n} = \frac{c_n}{\nu_n} f_1 \circ \frac{c_1}{\nu_1^{\eta_1+1}} f_2 \circ \dots \circ \frac{c_{n-1}}{\nu_{n-1}^{\eta_{n-1}+1}} f_n \left(\frac{1}{\nu_n^{\eta_n}} (x^*)^{n,\eta_n} \right).$$

Since $\frac{c_n}{\nu_n} f_1 \circ \frac{c_1}{\nu_1^{\eta_1+1}} f_2 \circ \dots \circ \frac{c_{n-1}}{\nu_{n-1}^{\eta_{n-1}+1}} f_n \left(\frac{1}{\nu_n^{\eta_n}} \cdot \right)$ is decreasing, there exists exactly one solution $(x^*)^{n,\eta_n}$ in \mathbb{R} . Once $(x^*)^{n,\eta_n}$ is fixed, we obviously have $(x^*)^{n,\eta_{n-1}} = \frac{1}{\nu_n} (x^*)^{n,\eta_n}$, and the values of the other coordinates of x^* follow in a similar way. •

In special cases system (4.19) is necessarily attracted to a non-equilibrium periodic orbit. Let $\kappa := n + \sum_{k=1}^n \eta_k$ be the dimension of (4.19).

We introduce the following assumption.

Assumption 4 *Suppose that $f_k, 1 \leq k \leq n$, are non-decreasing bounded analytic functions. Moreover, suppose that $\rho := \prod_{k=1}^n c_k f'_k((x^*)^{k,0})$ satisfies that $\rho < 0$.*

Notice that under Assumption 4, the conditions of Proposition 4 are satisfied, and thus (4.19) admits a unique equilibrium x^* under Assumption 4.

The following theorem is based on Theorem 4.3 of [30] and generalizes the result obtained by Theorem 6.3 in [2].

Theorem 3 *Grant Assumption 4. Consider all solutions λ of*

$$(\nu_1 + \lambda)^{\eta_1+1} \cdot \dots \cdot (\nu_n + \lambda)^{\eta_n+1} = \rho \quad (4.20)$$

and suppose that there exist at least two solutions λ of (4.20) such that

$$\operatorname{Re}(\lambda) > 0. \quad (4.21)$$

(i) x^ is linearly unstable, and the system (4.19) possesses at least one, but no more than a finite number of periodic orbits. At least one of them is orbitally asymptotically stable.*

(ii) Moreover, if $\kappa = 3$, then there exists a globally attracting invariant surface Σ such that x^ is a repeller for the flow in Σ . Every solution of (4.19) will be attracted to a non constant periodic orbit.*

Proof Since all functions f_k are bounded, the system (4.19) possesses a compact invariant set K . Rewriting (4.19) as $\dot{x} = F(x)$, where $x = (x^{1,0}, \dots, x^{n,\eta_n})^T$, the characteristic polynomial $P(\lambda)$ of $DF(x^*)$ is given by

$$P(\lambda) = \prod_{k=1}^n (-\nu_k - \lambda)^{\eta_k+1} - (-1)^\kappa \rho = (-1)^\kappa \left[\prod_{k=1}^n (\nu_k + \lambda)^{\eta_k+1} - \rho \right].$$

By assumption, there exist at least two eigenvalues having strictly positive real part. Therefore x^* is unstable. Moreover, since $\rho < 0$,

$$\det(-DF(x^*)) > 0$$

which is condition (4.5) of [30]. Then, the last assertion of item (i) follows from Theorem 4.3 of [30].

To prove part (ii), we follow the proof of Theorem 6.3 in [2]. First, notice that $DF(x^*)$ is given by the matrix

$$\begin{pmatrix} -d & a & 0 \\ 0 & -e & b \\ c & 0 & -f \end{pmatrix},$$

where $d, e, f > 0$ and $\rho = abc < 0$. Hence, either all a, b, c are negative or only one of them, say c , is negative. In the first case, $-DF(x^*)$ is a positive irreducible matrix, in the second case, the change of variables $y_1 = x_1, y_2 = -x_2, y_3 = x_3$ leads to a negative irreducible matrix. We therefore suppose without loss of generality that we are in the first case. Then the Perron-Frobenius theorem implies that $DF(x^*)$ possesses a single largest eigenvalue which is strictly negative, and the eigenvector associated to it has all its components of the same sign. Moreover, the other two eigenvectors associated to the conjugate complex eigenvalues having positive real part do not have all components of the same sign.

By Theorem 1.7 of Hirsch (1988) [23], there exists a globally attracting invariant surface Σ such that every trajectory within the invariant set K is eventually attracted to Σ . By [2] the equilibrium x^* is a repeller for the flow in K . Hence, the Poincaré-Bendixson theorem implies that each such trajectory will eventually converge to a non constant periodic orbit.

•

Remark 3 If $\nu_1 = \dots = \nu_n = \nu$, then for $\kappa \geq 3$ the following condition

$$|\rho| > \frac{\nu^\kappa}{\left(\cos\left(\frac{\pi}{\kappa}\right)\right)^\kappa} \quad (4.22)$$

implies (4.21). Indeed, the different eigenvalues for $1 \leq j \leq \kappa$ are given by

$$\lambda_j = -\nu - |\rho|^{\frac{1}{\kappa}} e^{i\frac{2j\pi}{\kappa}} \quad (\kappa \text{ odd}) \quad ; \quad \lambda_j = -\nu + |\rho|^{\frac{1}{\kappa}} e^{i\frac{(2j-1)\pi}{\kappa}} \quad (\kappa \text{ even}).$$

If κ is odd, then there is exactly one real root, which is strictly negative, $\lambda_\kappa = -\nu - |\rho|^{\frac{1}{\kappa}}$. The rest are complex conjugate pairs with real part $-\nu - |\rho|^{\frac{1}{\kappa}} \cos\left(\frac{2j\pi}{\kappa}\right)$. The maximal value is $-\nu + |\rho|^{\frac{1}{\kappa}} \cos\left(\frac{\pi}{\kappa}\right)$ for $j = (\kappa \pm 1)/2$, such that (4.22) implies (4.21). If κ is even, then all roots are complex conjugate pairs with real part $-\nu + |\rho|^{\frac{1}{\kappa}} \cos\left(\frac{(2j-1)\pi}{\kappa}\right)$. The maximal value is as before, now for $j = 1, \kappa$, such that again (4.22) implies (4.21).

Remark 4 (Phase transition due to increasing memory) *In some cases, increasing the order of the memory, i.e. the value of some of the exponents η_k in (4.17) or equivalently the value of κ , can lead to a phase transition within the system (4.19). At the phase transition point, a system which was stable can become unstable, and in certain cases, increasing the order even more might stabilize the system again. As an example, consider a family of n populations of neurons, where $n > 1$ is fixed, and such that $\nu_k = \nu$ for all $1 \leq k \leq n$. If $\kappa = 2$, the fixed point is stable since eigenvalues are $\lambda_j = -\nu \pm i\sqrt{|\rho|}$, and only damped oscillations occur. We will assume $\kappa \geq 3$.*

First note that ρ is bounded due to the Lipschitz condition on the rate functions f_k . The right hand side of (4.22) goes to infinity for $\nu \rightarrow \infty$ for all values of κ , and thus, if ν is large, the system will always be stable and not exhibit oscillations. For any fixed value of $\nu > 1$, it also goes to infinity for $\kappa \rightarrow \infty$, such that a possible unstable system becomes stable for increasing κ . This implies that for a discrete delay of any value the system will never exhibit oscillations, since a discrete delay is obtained for $\eta_k \rightarrow \infty$, keeping η_k/ν_k constant.

Now assume that $\nu = 1$. Then increasing κ does not change the coordinates $(x^)^{k,0}$ of the equilibrium state x^* , so ρ does not change. The right hand side of (4.22) decreases towards one, so if $-8 < \rho < -1$, then there exists $\kappa_0 > 3$ minimal such that for all $\kappa \geq \kappa_0$, (4.22) is fulfilled, but $|\rho| \leq (\nu/\cos(\frac{\pi}{\kappa}))^\kappa$ for $\kappa < \kappa_0$. Then all models corresponding to $\kappa < \kappa_0$ have x^* as attracting equilibrium point, but for $\kappa \geq \kappa_0$, the equilibrium x^* becomes unstable.*

As a corollary of the above Theorem, we show that one of the conditions needed to state the central limit theorem in Theorem 2 is satisfied.

Corollary 1 *Suppose that $n = 2$ and that the conditions of Theorem 3 hold true. Then there exist $\alpha_1, \alpha_2 > 0$ such that*

$$\liminf_{t \rightarrow \infty} \frac{m_t^k}{t} \geq \alpha_k, \quad k = 1, 2.$$

Proof Since $c_1 c_2 < 0$, any solution to (4.19) is eventually attracted to a non constant periodic orbit. Since $m_t^k = \int_0^t f_k(x_s^{k,0}) ds$ and since f_k is non-decreasing and strictly positive, it follows that $\liminf_{t \rightarrow \infty} \frac{1}{t} m_t^k > 0$. •

5 Study of an approximating diffusion process and simulation study

In this section we work with the cyclic feedback system of the last section. The aim is to study to which extent the behavior of the limit system is also observed within the finite size system $Z_{k,i}^N$.

5.1 An associated system of piecewise deterministic Markov processes

Introducing the family of adapted càdlàg processes (recall (4.16))

$$X_k^N(t) := \frac{1}{N_{k+1}} \sum_{j=1}^{N_{k+1}} \int_{]0,t]} h_{kk+1}(t-s) dZ_{k+1,j}^N(s) = \int_{]0,t]} h_{kk+1}(t-s) d\bar{Z}_{k+1}^N(s), \quad (5.23)$$

where $\bar{Z}_{k+1}^N(s) = \frac{1}{N_{k+1}} \sum_{j=1}^{N_{k+1}} Z_{k+1,j}^N(s)$ and recalling (2.3), it is clear that the dynamics of the system is entirely determined by the dynamics of the processes $X_k^N(t-), t \geq 0$.¹ In some sense, X_k^N describes the accumulated memory belonging to the directed edge pointing from population $k+1$ to population k . Without assuming the memory kernels to be Erlang kernels, the system $(X_k^N, 1 \leq k \leq n)$ is not Markovian. For general memory kernels, Hawkes processes are truly infinite memory processes.

When the kernels are Erlang, given by (4.17), taking formal derivatives in (5.23) with respect to time t and introducing for any k and $0 \leq l \leq \eta_k$

$$X_{k,l}^N(t) := c_k \int_{]0,t]} \frac{(t-s)^{\eta_k-l}}{(\eta_k-l)!} e^{-\nu_k(t-s)} d\bar{Z}_{k+1}^N(s), \quad (5.24)$$

we obtain the following system of stochastic differential equations which is a stochastic version of (4.19).

$$\begin{cases} dX_{k,l}^N(t) &= [-\nu_k X_{k,l}^N(t) + X_{k,l+1}^N(t)]dt, \quad 0 \leq l < \eta_k, \\ dX_{k,\eta_k}^N(t) &= -\nu_k X_{k,\eta_k}^N(t)dt + c_k d\bar{Z}_{k+1}^N(t). \end{cases} \quad (5.25)$$

Here, X_k^N is identified with $X_{k,0}^N$, $\bar{Z}_k^N = \frac{1}{N_k} \sum_{j=1}^{N_k} Z_{k,j}^N$, and each $Z_{k,j}^N$ jumps at rate $f_k(X_{k,0}^N(t-))$. We call the system (5.25) a *cascade of memory terms*. Thus, the dynamics of the Hawkes process $(Z_{k,i}^N(t))_{(k,i) \in I^N}$ is entirely determined by the piecewise deterministic Markov process (PDMP) $(X_{k,l}^N)_{(1 \leq k \leq n, 0 \leq l \leq \eta_k)}$ of dimension κ .

5.2 A diffusion approximation in the large population regime

The process $\bar{Z}_{k+1}^N(t)$ appearing in the last equation of (5.25) jumps at a rate given by $N_{k+1}f_{k+1}(X_{k+1,0}^N(t-))$, having jumps of size $\frac{1}{N_{k+1}}$. Its variance is $\frac{f_{k+1}(X_{k+1,0}^N(t-))}{N_{k+1}}$. Therefore, it is natural to consider the approximating diffusion process

$$\begin{cases} dY_{k,l}^N(t) &= [-\nu_k Y_{k,l}^N(t) + Y_{k,l+1}^N(t)]dt, \quad 0 \leq l < \eta_k, \\ dY_{k,\eta_k}^N(t) &= -\nu_k Y_{k,\eta_k}^N(t)dt + c_k f_{k+1}(Y_{k+1,0}^N(t))dt + c_k \frac{\sqrt{f_{k+1}(Y_{k+1,0}^N(t))}}{\sqrt{N_{k+1}}} dB_{k+1}(t), \end{cases} \quad (5.26)$$

where the $B_l(t), 1 \leq l \leq n$, are independent standard Brownian motions, approximating the jump noise of each population. Write A^X for the infinitesimal generator of the process (5.25) and A^Y for the corresponding generator of (5.26). Moreover, write P_t^X and P_t^Y for the associated Markovian semigroups. We denote generic elements of the state space \mathbb{R}^κ of Y^N by $x = (x^1, \dots, x^\kappa)$. Finally, for a function g defined on \mathbb{R}^κ , we define

$$\|g\|_{r,\infty} := \sum_{k=0}^r \sum_{|\alpha|=k} \|\partial^\alpha g\|_\infty.$$

¹We have to take the left-continuous version, since intensities are predictable processes.

Then we obtain the following approximation result showing that Y^N is a good small noise approximation of X^N .

Theorem 4 *Suppose that all spiking rate functions f_k belong to the space C_b^5 of bounded functions having bounded derivatives up to order 5. Then there exists a constant C depending only on f_1, \dots, f_n and the bounds on its derivatives such that for all $\varphi \in C_b^4(\mathbb{R}^\kappa; \mathbb{R})$,*

$$\|P_t^X \varphi - P_t^Y \varphi\|_\infty \leq Ct \frac{\|\varphi\|_{4,\infty}}{N^2}.$$

The proof is given in the Appendix.

Theorem 4 is a first step towards convergence in law and shows that the diffusion process (5.26) is a good approximation of (5.25), as $N \rightarrow \infty$. However, in the limit of $N \rightarrow \infty$, (5.26) is not a diffusion anymore, since the diffusive term tends to zero. Both processes, X^N and Y^N , tend to the limit process described in section 2.2. This convergence is of rate $\frac{1}{N}$, which is slower than the approximation proved in Theorem 4.

5.3 Oscillations of the approximating diffusion at fixed population size

We now show to which extent the approximating diffusion process (5.26) imitates the oscillatory behavior of the limit system described in section 4. Consider two populations, $n = 2$, where the memory kernels are given by (4.17).

We denote by

$$b(x) := \begin{pmatrix} -\nu_1 x^1 + x^2 \\ -\nu_1 x^2 + x^3 \\ \vdots \\ -\nu_1 x^{\eta_1+1} + c_1 f_2(x^{\eta_1+2}) \\ -\nu_2 x^{\eta_1+2} + x^{\eta_1+3} \\ \vdots \\ -\nu_2 x^\kappa + c_2 f_1(x^1) \end{pmatrix} \quad (5.27)$$

the drift vector of (5.26). Moreover, we introduce the $\kappa \times 2$ -diffusion matrix

$$\sigma(x) := \begin{pmatrix} 0 & 0 \\ \vdots & \vdots \\ 0 & \frac{c_1}{\sqrt{p_2}} \sqrt{f_2(x^{\eta_1+2})} \\ 0 & 0 \\ \vdots & \vdots \\ \frac{c_2}{\sqrt{p_1}} \sqrt{f_1(x^1)} & 0 \end{pmatrix}, \quad (5.28)$$

where $p_1 = N_1/N, p_2 = N_2/N$. Then we may rewrite (5.26) as

$$dY^N(t) = b(Y^N(t))dt + \frac{1}{\sqrt{N}} \sigma(Y^N(t))dB(t), \quad (5.29)$$

with $B(t) = (B^1(t), B^2(t))^T$.

Throughout this section, we assume the conditions of Theorem 3, in particular, suppose that $c_1 c_2 < 0$. Moreover, suppose that f_1 and f_2 are smooth strictly positive non-decreasing functions. In this case, under condition (4.21), the associated limit system possesses a non constant periodic orbit which is asymptotically orbitally stable. We will now show that also the finite size system (5.29) is attracted to this periodic orbit.

The existence of a global Lyapunov function implies that there exists a compact set K such that process (5.29) visits K infinitely often, almost surely. More precisely, recalling that A^Y denotes the infinitesimal generator of (5.29), we have the following result. In order to simplify notation, in the next proposition, we write $x = (x^{1,0}, \dots, x^{1,\eta_1}, x^{2,0}, \dots, x^{2,\eta_2})$ for generic elements of \mathbb{R}^κ .

Proposition 5 *Grant the conditions of Theorem 3. Let $j(x)$ be a smoothed version of $|x|$, i.e., $j(x) = |x|$ for all $|x| \geq 1$, $|j'(x)| \leq C$, $|j''(x)| \leq C$, for some constant C , for all x . Put then $G(x) := \sum_{k=1}^2 \sum_{l=0}^{\eta_k} \frac{l+1}{\nu_k^l} j(x^{k,l})$. Then G is a Lyapunov-function for (5.29) in the sense that*

$$A^Y G(x) \leq -cG(x) + d,$$

for some constants $c, d > 0$ depending on $\max_k \|f_k\|_\infty$.

Proof The above statement follows from the fact that the drift part of $A^Y G(x)$ is given by

$$\sum_{k=1}^2 \sum_{j=0}^{\eta_k-1} \left\{ -\nu_k x^{k,j} + x^{k,j+1} \right\} \frac{\partial G}{\partial x^{k,j}} - \sum_{k=1}^2 \nu_k x^{k,\eta_k} \frac{\partial G}{\partial x^{k,\eta_k}} + \sum_{k=1}^2 c_k f_{k+1}(x^{k+1,0}) \frac{\partial G}{\partial x^{k,\eta_k}}.$$

But for $|x^{k,j}| \geq 1$,

$$\begin{aligned} \left\{ -\nu_k x^{k,j} + x^{k,j+1} \right\} \frac{\partial G}{\partial x^{k,j}} &= \left\{ -\nu_k x^{k,j} + x^{k,j+1} \right\} \text{sign}(x^{k,j}) \frac{(j+1)}{(\nu_k)^j} \\ &\leq -\frac{(j+1)}{\nu_k^{j-1}} |x^{k,j}| + \frac{(j+1)}{\nu_k^j} |x^{k,j+1}|. \end{aligned}$$

As a consequence, if $|x^{k,j}| \geq 1$ for all k, j , the contribution of the drift part can be upper bounded by

$$-\sum_{k=1}^2 \nu_k |x^{k,0}| - \sum_{k=1}^2 \sum_{j=1}^{\eta_k} \frac{1}{\nu_k^{j-1}} |x^{k,j}| + 2 \max\left(\frac{\eta_1+1}{\nu_1^{\eta_1}} \|f_1\|_\infty, \frac{\eta_2+1}{\nu_2^{\eta_2}} \|f_2\|_\infty\right) \leq -cG(x) + d,$$

for some constants $c, d > 0$. On the other hand, the contributions coming from terms with $|x^{k,j}| \leq 1$ are bounded, and the contribution coming from the diffusion part of $A^Y G$ is bounded as well. This finishes the proof. \bullet

In particular, putting $K = \{G \leq 2d/c\}$, it follows that

$$A^Y G(x) \leq -\frac{c}{2} G(x) + d \mathbf{1}_K(x). \quad (5.30)$$

It is well-known (see e.g. Douc, Fort and Guillin (2009) [15]) that (5.30) implies that

$$\mathbb{E}_x(e^{\frac{c}{2} \tau_K}) \leq G(x), \quad (5.31)$$

where $\tau_K = \inf\{t \geq 0 : Y^N(t) \in K\}$. Thus, the process comes back to the compact K infinitely often, and excursions out of K have exponential moments. In particular, we can concentrate on the study of the trajectories inside K .

We will now study the behavior of the trajectories of Y^N inside the compact K . By Theorem 3, under condition (4.21), the limit system possesses a non constant periodic orbit which is asymptotically orbitally stable. Denote this orbit by Γ and let T be its periodicity. We suppose without loss of generality that $\Gamma \subset K$. In the following, we will show that each time the process is inside K , it will also visit vicinities of the periodic orbit (in a sense that will be made precise in Theorem 5 below). We start with support properties.

Fix $\varepsilon > 0$ and let $S(\varepsilon, \Gamma) := \{x : d(x, \Gamma) < \varepsilon\}$ be a tube around this orbit. Denote by Q_x^Y the law of the solution $(Y^N(t), t \geq 0)$ of (5.29), starting from $Y^N(0) = x$. Fix $t_1 > 1$ and let $O = \{\varphi \in C(\mathbb{R}_+; \mathbb{R}^\kappa) : \varphi(t) \in S(\varepsilon, \Gamma) \forall 1 \leq t \leq t_1\}$. Then we have the following first result concerning the support of Q_x^Y .

Proposition 6 *Under the conditions of Theorem 3 and supposing that f_1 and f_2 are strictly positive, the following holds true. For all $x \in \mathbb{R}^\kappa$, we have that $Q_x^Y(O) > 0$.*

Hence, the diffusion Y^N visits the tube around Γ , starting from any initial point $x \in \mathbb{R}^\kappa$, during the time interval $[1, t_1]$, with positive probability. The fact that we choose time intervals $[1, t_1]$ is not important, and we could equally work with any time interval $[t_0, t_1]$, for any fixed $t_0 > 0$, see the proof in Appendix 6.4.

Note that the above proposition gives a statement concerning the support of the law of Y^N for fixed N . Its proof relies on the support theorem for diffusions. The result does not say anything about the actual value of the probability of tubes around Γ , nor about the precise $N \rightarrow \infty$ -asymptotics : this is outside the scope of the present paper and will be the subject of a future work.

Proposition 6 implies that there is strictly positive probability to visit tubes around Γ , yet, we do not know that such visits arrive almost surely, within a finite time horizon. In order to prove this, we will show that $x \mapsto Q_x^N(O)$ is lower-bounded on compacts. Thus, we have to control the dependence on the starting configuration x in the above proposition. Therefore, we show that $x \mapsto Q_x^Y(A)$ is continuous for Borel sets $A \in \mathcal{B}(C(\mathbb{R}_+; \mathbb{R}^\kappa))$ which are e.g. of the form O , i.e., we need to show that the process Y^N is strong Feller. However, Y^N is a degenerate diffusion process since Brownian noise only appears in the coordinates Y_{1, η_1}^N and Y_{2, η_2}^N . Following Ishihara and Kunita (1974) [26], Lemma 5.1, we therefore show that the process satisfies the weak Hörmander condition. Notice that due to the specific form of the diffusion matrix in equation (5.28), the Itô and the Stratonovich forms are the same.

Recall that for smooth vector fields $f(x)$ and $g(x) : \mathbb{R}^\kappa \rightarrow \mathbb{R}^\kappa$, the Lie bracket $[f, g]$ is defined by

$$[f, g]^i = \sum_{j=1}^{\kappa} \left(f^j \frac{\partial g^i}{\partial x^j} - g^j \frac{\partial f^i}{\partial x^j} \right), \quad i = 1, \dots, \kappa.$$

We write $\sigma^1, \sigma^2 : \mathbb{R}^\kappa \rightarrow \mathbb{R}^\kappa$ for the two column vectors constituting the diffusion matrix σ of (5.28).

Definition 3 Define a set \mathcal{L} of vector fields by the ‘initial condition’ $\sigma^1, \sigma^2 \in \mathcal{L}$ and an arbitrary number of iteration steps

$$L \in \mathcal{L} \implies [b, L], [\sigma^1, L], [\sigma^2, L] \in \mathcal{L}. \quad (5.32)$$

For $M \in \mathbb{N}$, define the subset \mathcal{L}_M by the same initial condition and at most M iterations (5.32). Write \mathcal{L}_M^* for the closure of \mathcal{L}_M under Lie brackets; finally, write

$$\Delta_{\mathcal{L}_M^*} := \text{LA}(\mathcal{L}_M)$$

for the linear hull of \mathcal{L}_M^* , i.e. the Lie algebra spanned by \mathcal{L}_M .

Definition 4 We say that a point $z^* \in \mathbb{R}^\kappa$ is of full weak Hörmander dimension if there is some $M \in \mathbb{N}$ such that

$$(\dim \Delta_{\mathcal{L}_M^*})(z^*) = \kappa. \quad (5.33)$$

Due to the cascade structure of the drift vector b and since $f_1(\cdot), f_2(\cdot) > 0$ on \mathbb{R} , it is straightforward to show that the weak Hörmander dimension holds at all points $x \in \mathbb{R}^\kappa$:

Proposition 7 Suppose that f_1 and f_2 are smooth and strictly positive. Then for all $x \in \mathbb{R}^\kappa$, $(\dim \Delta_{\mathcal{L}_M^*})(x) = \kappa$, where $M = \max(\eta_1, \eta_2)$.

Proof The proof is done by first calculating the Lie-bracket $[\sigma^1, b]$ and $[\sigma^2, b]$ and then successively bracketing with b . •

Once the weak Hörmander condition holds everywhere, it follows that the process is strongly Feller, see [26]. As a consequence, the following holds.

Corollary 2 Let $A \in \mathcal{B}(C(\mathbb{R}_+; \mathbb{R}^\kappa))$ be of the type $A = \{\varphi : \varphi(t) \in B \ \forall t_1 \leq t \leq t_2\}$, for some $B \in \mathcal{B}(\mathbb{R}^\kappa)$ and for $t_1 < t_2$. Then

$$x \mapsto Q_x^Y(A) \text{ is continuous.}$$

Proof Using the Markov property, we have

$$Q_x^Y(A) = \int P_{t_1}^Y(x, dy) Q_y^Y(\{\varphi : \varphi(t) \in B \ \forall 0 \leq t \leq t_2 - t_1\}) = P_{t_1}^Y \Phi(x),$$

where $\Phi(y) = Q_y^Y(\{\varphi : \varphi(t) \in B \ \forall 0 \leq t \leq t_2 - t_1\})$ is measurable and bounded. But $P_{t_1}^Y \Phi(x)$ is continuous in x , since Y^N is strong Feller. •

A direct consequence of the above result is the fact that

$$K \ni x \mapsto Q_x^Y(O)$$

is strictly lower bounded for any fixed $t_1 > 1$. This fact enables us to conclude our discussion and to show that the diffusion approximation (5.29) will have the same type of oscillations as the limit system (m_t^1, m_t^2) . Let

$$\tau_\Gamma(t_1) := \inf\{t \geq 0 : Y^N(s) \in S(\varepsilon, \Gamma) \ \forall t \leq s \leq t + t_1\}.$$

Theorem 5 *Grant the assumptions of Theorem 3 and let Γ be a non constant periodic orbit of period T of the limit system which is asymptotically orbitally stable. Then for all $\varepsilon > 0$ and $t_1 > 1$, there exist $C, \lambda > 0$ such that*

$$\mathbb{E}_x(e^{\lambda \tau_\Gamma(t_1)}) \leq CG(x).$$

Moreover,

$$\limsup_{t \rightarrow \infty} \mathbf{1}_{\{Y^N(s) \in S(\varepsilon, \Gamma) \ \forall t \leq s \leq t+t_1\}} = 1$$

\mathbb{P}_x -almost surely, for all $x \in \mathbb{R}^\kappa$.

In particular, the above theorem implies that the process Y^N visits the oscillatory region $S(\varepsilon, \Gamma)$ during time intervals of length t_1 infinitely often. The choice of t_1 in the above theorem is free. By choosing t_1 larger than the period T of Γ (or even choosing $t_1 \geq kT$ for some fixed $k \in \mathbb{N}$), this implies that Y^N will present oscillations infinitely often almost surely.

Proof By Proposition 6, for every x and every fixed $t_1 > 1$, $Q_x^Y(O) > 0$. By continuity and since K is compact, this implies that

$$\inf_{x \in K} Q_x^Y(O) > 0.$$

Now, by (5.31), the process visits the compact set K infinitely often, almost surely, with exponential moments for the successive visits of the process to K . The assertion then follows by the conditional version of the Borel-Cantelli lemma. \bullet

Remark 5 *The above result shows that Y^N does actually visit tubes around the periodic orbit infinitely often, with waiting times that possess exponential moments. It is valid for a fixed population size N . The above result does not show that, starting from a vicinity of Γ , the diffusion stays there for a long time, before being kicked out of the tube due to noise. Such a study is much more difficult, it is related to the large deviation properties of the process and will be the topic of a future work.*

5.4 Simulation study

In this Section we will check by simulations how N influences the approximating diffusion, and compare the behavior of (4.19) and (5.26). We set $n = 2$ and $c_1 = -1, c_2 = 1$ such that population 1 is inhibitory and population 2 is excitatory, and thus $\rho < 0$, and it is a negative feedback system. We will use the following bounded, Lipschitz and strictly increasing rate functions,

$$f_1(x) = \begin{cases} 10e^x & \text{for } x < \log(20) \\ \frac{400}{1+400e^{-2x}} & \text{for } x \geq \log(20) \end{cases} \quad ; \quad f_2(x) = \begin{cases} e^x & \text{for } x < \log(20) \\ \frac{40}{1+400e^{-2x}} & \text{for } x \geq \log(20) \end{cases}.$$

In the first set of simulations, we put $\nu_1 = \nu_2 = 1$ and $\eta_1 = 3, \eta_2 = 2$ such that $\kappa = 7$. Then $(x^*)^{1,l} = -2.424$ for $l = 0, 1, \dots, \eta_1$, and $(x^*)^{2,l} = 0.885$ for $l = 0, 1, \dots, \eta_2$. This yields $\rho = -2.15$ and $\nu^\kappa / (\cos(\frac{\pi}{\kappa}))^\kappa = 2.08$, and thus, (4.22) is fulfilled. The period is approximately $2\pi/\omega = 12.98$, where $\omega = |\rho|^{1/\kappa} \sin(\pi/\kappa)$. Finally, we put $N_1 = N_2 = 20$. Results are presented in Fig. 1. The cascade structure in the memory variables is clearly

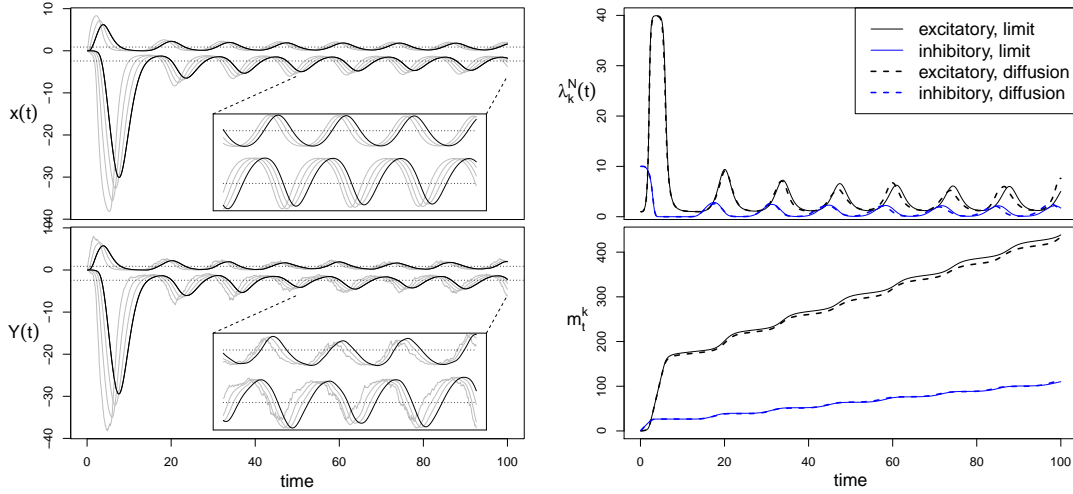


Figure 1: Comparison of the limit system and the diffusion approximation through simulations. Left top panel: The limiting system (4.19). The black curves are $x^{k,0}$ given in (4.16), and those that are felt by the system. The gray curves are the variables in the Markovian memory cascade. The dotted lines indicate the location of the critical point. The inset is a blow-up of the last part of the simulation. Parameters are $c_1 = -1, c_2 = 1, \nu_1 = \nu_2 = 1, \eta_1 = 3, \eta_2 = 2$. Left bottom panel: As above but for the approximating system (5.26), with $N_1 = N_2 = 20$. Right top panel: Intensity processes corresponding to the simulations on the left. Right bottom panel: The cumulative intensity processes. As predicted by the Central Limit Theorem 2, both approximations are getting worse as t increases.

seen, and the noisy diffusion approximation follows the limit cycle. The periodic behavior is evident.

To investigate how close the approximating diffusion follows the oscillatory behavior of the limit system, we simulated 20 repetitions of the noisy process for different values of $N = 20, 100, 200, 1000$ and for $p_1 = p_2 = 1/2$, and compared it to the limit system on a later time interval in Fig. 2. For small N , the system shifts phase randomly relative to the limiting system, but it maintains the oscillations. For larger N , the system follows the limiting system closely on a longer time horizon.

To study phase transitions, we put $\nu = \nu_1 = \nu_2 = 0.8$ and for $\eta_1 = \eta_2$ we varied $\kappa = 4, 8, 12, 16, 20, 24$. Condition (4.22) is only fulfilled for $6 \leq \kappa \leq 12$. Results are in Fig. 3, left. For $\kappa = 4$ the condition is not fulfilled, and damped oscillations are seen. Then increasing κ , a phase transition occurs, yielding sustained oscillations. A further phase transition occurs when κ becomes larger than 12. For values of κ between 12 and 16, damped oscillations happen after the initial large excursion, and when κ becomes even larger, the system converges to the steady state in a seemingly monotone manner. Note that for $\nu \neq 1$ the steady state depends on the order of the memory.

Finally, we fix $\eta_1 = \eta_2 = 3$ and vary ν in Fig. 3, right. For low and high values of ν , no oscillations occur, but at $\nu_{low} = 0.7169$ a Hopf-bifurcation occurs, and another again at $\nu_{high} = 1.3982$. Thus, the system oscillates in an interval $\nu \in (\nu_{low}, \nu_{high})$. In general, the interval will depend on the order of the memory.

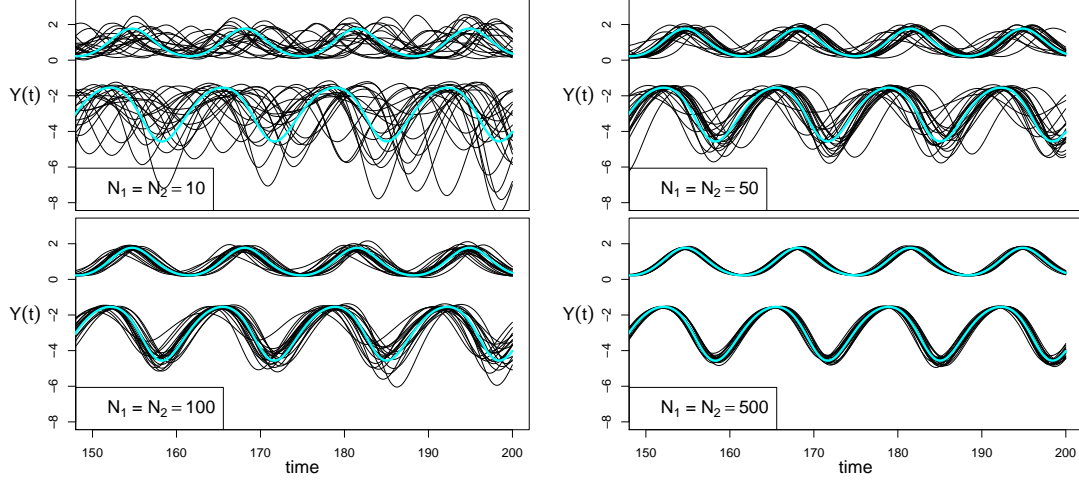


Figure 2: Diffusion approximation for increasing N , each panel contains 20 realizations. The cyan curve is the limit system. Parameters are as in Fig. 1.

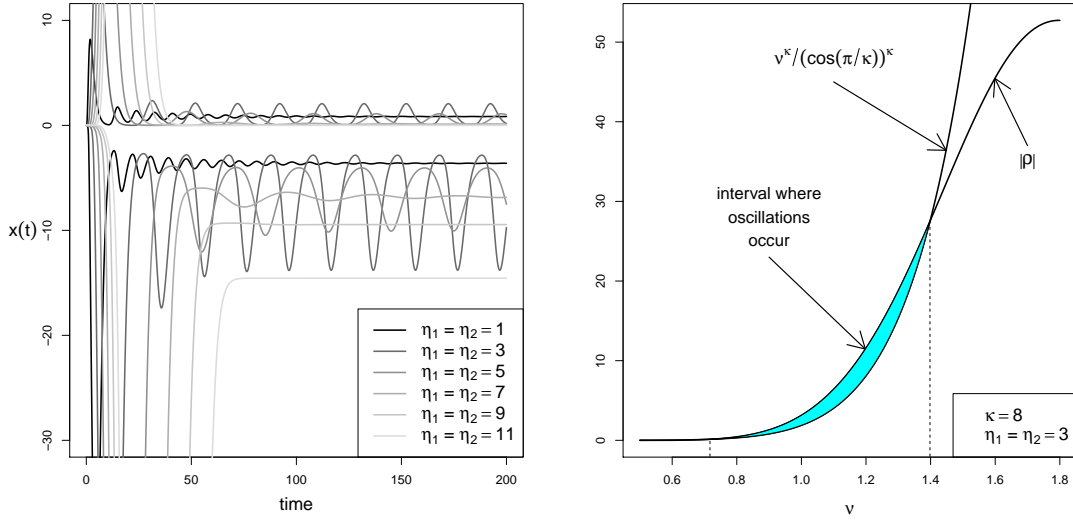


Figure 3: Phase transitions and Hopf bifurcations. Left: Phase transitions due to increasing order of the memory. Parameters are $\nu_1 = \nu_2 = 0.8$ and varying $\eta_1 = \eta_2$, so that $\kappa = 2(1+\eta_1)$. Right: Check of condition (4.22) as a function of ν . In the cyan area, at least two eigenvalues of $DF(x^*)$ have positive real part, and self-sustained oscillations occur in system (4.19). Hopf bifurcations happen at values $\nu_{low} = 0.7169$ and at $\nu_{high} = 1.3982$.

6 Appendix

In the proofs below, C denotes a generic constant, which might change value from equation to equation, and from line to line, even within the same equation.

6.1 Convolution equations in the matrix case

The following is a version of Lemma 26 of [13] in the multidimensional case. It is based on old results of [1, 11] on systems of renewal equations.

Lemma 1 (Corollary 3.1 of [11], see also Theorem 2.2 of [1]) *Let $H(s)$ be given by (3.13) such that Assumption 3 is fulfilled (supercritical case). Put $\Gamma(t) = \sum_{n \geq 1} H^{*n}(t)$. Then the following assertions hold.*

1. *There is a unique α_0 such that $\int_0^\infty e^{-\alpha_0 t} H(t) dt$ has largest eigenvalue $\equiv +1$.*
2. *Γ is locally bounded. Moreover, for some constant C ,*

$$\Gamma_{ij}(t) \leq C e^{\alpha_0 t}. \quad (6.34)$$

3. *For any pair of locally bounded functions $u, h : \mathbb{R}_+ \rightarrow \mathbb{R}^n$ such that $u = h + H * u$, it holds $u = h + \Gamma * h$.*

Proof 1. Due to Assumption 3, there exist constants C and p such that $|h_{kl}|(t) \leq C(1 + t^p)$, for all $1 \leq k, l \leq n$, for all $t \geq 0$. As a consequence, the matrix-Laplace transform

$$\mathcal{L}_H(\alpha) := \int_0^\infty e^{-\alpha t} H(t) dt$$

is well-defined for all $\alpha > 0$. Being a primitive matrix, by the Perron-Frobenius theorem, it possesses a unique maximal eigenvalue $\lambda_H(\alpha)$ with an associated eigenvector composed of positive coordinates. By Assumption 3, $\lambda_H(0) > 1$ and $\lambda_H(\infty) = 0$. This implies that there exists a unique α_0 such that $\lambda_H(\alpha_0) = 1$. (This step demands some extra work, which is done in [1, 11]).

2. Let $M(t) = e^{-\alpha_0 t} H(t)$. Then $M * M(t) = e^{-\alpha_0 t} H * H(t)$, thus, $\sum_{n \geq 1} M^{*n}(t) = e^{-\alpha_0 t} \Gamma(t)$. Since $H(s)$ is at most of polynomial growth, then $\int_0^\infty t M(t) dt < \infty$ component-wise. This implies that all entries of the matrix \mathcal{H} of (3.9) of [11] are finite, so that $\lim_{t \rightarrow \infty} \sum_{n \geq 1} (M^{*n}(t))_{i,j} / t$ exists and is finite, for all $1 \leq i, j \leq n$, see page 431 of the proof of Theorem 3.1 of [11], see also Theorem 2.2 of [1]. It follows that there exists a constant C such that $\Gamma_{i,j}(t) \leq C e^{\alpha_0 t}$, for all $1 \leq i, j \leq n$.

3. Follows from Theorem 2.1 of [11]. •

6.2 Proof of Theorem 2

We first prove Propositions 2 and 3.

Proof of Proposition 2 We have

$$\lambda_t^k := \frac{dm_t^k}{dt} = f_k \left(\int_0^t \sum_l h_{kl}(t-s) dm_s^l \right).$$

Using the Lipschitz property of f_k , we obtain

$$\lambda_t^k \leq f_k(0) + \sum_{l=1}^n L \int_0^t |h_{kl}(t-s)| \lambda_s^l ds. \quad (6.35)$$

Using $H(s)$ defined in (3.13), we can rewrite (6.35) as

$$\lambda_t \leq f(0) + H * \lambda(t),$$

where $f(0) = (f_1(0), \dots, f_n(0))^T$. Define $\Gamma_l := \sum_{1 \leq k \leq l} H^{*k}$. Since for any two matrix-valued functions A, B , $\int_0^\infty A * B(t) dt = (\int_0^\infty A(t) dt)(\int_0^\infty B(t) dt)$, provided the integrals are well-defined, we have

$$\int_0^\infty \Gamma_l(t) dt = \sum_{1 \leq k \leq l} \Lambda^k, \quad \Lambda = \int_0^\infty H(t) dt,$$

having unique maximal eigenvalue $\sum_{1 \leq k \leq l} \mu_1^k \leq \frac{\mu_1}{1-\mu_1}$. As a consequence, the renewal function $\Gamma(t) = \sum_{l \geq 1} H^{*l}(t)$ is well-defined and locally bounded, and the maximal eigenvalue of $\int_0^\infty \Gamma(t) dt$ is given by $\frac{\mu_1}{1-\mu_1}$.

Any solution $a(t)$ of $a = f(0) + H * a$ is given by $a(t) = f(0) + \Gamma * f(0) = f(0) + \int_0^t \Gamma(s) f(0) ds$. Therefore,

$$\lambda(t) \leq f(0) + \left(\int_0^t \Gamma(s) ds \right) f(0) \leq f(0) + \left(\int_0^\infty \Gamma(s) ds \right) f(0),$$

where the inequality has to be understood component-wise. Thus, $\lambda(t)$ is a bounded function of t . This implies the first result.

By (2.9) and (2.10), writing $\delta^N(t) = (\delta_1^N(t), \dots, \delta_n^N(t))^T$,

$$\begin{aligned} \delta_k^N(t) &\leq (H * \delta^N)_k(t) + \frac{C}{\sqrt{N}} \int_0^t \sum_l \left[\int_0^s h_{kl}^2(s-u) \lambda_u^l du \right]^{1/2} ds \\ &\leq (H * \delta^N)_k(t) + \frac{Ct}{\sqrt{N}}, \end{aligned}$$

by the first step, since λ_u^l are bounded, and since h_{kl} by assumption are in $L^2(\mathbb{R}_+; \mathbb{R})$. As a consequence,

$$\delta_k^N(t) \leq \frac{Ct}{\sqrt{N}} + \frac{C}{\sqrt{N}} \int_0^t \sum_{l=1}^n \Gamma_{kl}(t-s) s ds \leq \frac{Ct}{\sqrt{N}}.$$

Together with the proof of Theorem 1 this finishes the proof. •

Proof of Proposition 3 As in the proof of Proposition 2, we obtain (6.35) for $1 \leq k \leq n$. Hence, using Lemma 1,

$$\lambda_t^k \leq f_k(0) + \sum_l \left(\int_0^t \Gamma_{kl}(s) ds \right) f_l(0) \leq f_k(0) + C e^{\alpha_0 t} \leq c e^{\alpha_0 t},$$

where $c = \max_{1 \leq k \leq n} f_k(0) + C$. The second assertion follows then analogously to the proof given for Proposition 2. •

The main ingredients for the proof of Theorem 2 are the bounds in Propositions 2 and 3, depending on the criticality. We now prove Theorem 2 in the subcritical case, which is an adaptation of the proof of Theorem 10 of [13] to the nonlinear case.

Proof of Theorem 2, subcritical case. Put $U_{k,i}^N(t) := Z_{k,i}^N(t) - m_t^k$, $(k, i) \in I^N$, and introduce the martingales

$$M_{k,i}^N(t) = \int_0^t \int_0^\infty \mathbf{1}_{\{z \leq f_k(\int_0^s h_{kl}(s-u) dm_l(u))\}} \tilde{N}_{k,i}^N(ds, dz) = \bar{Z}_{k,i}^N(t) - m_t^k,$$

where $\tilde{N}_{k,i}^N$ denotes the compensated PRM $N_{k,i}^N(ds, dz) - ds dz$. Then

$$U_{k,i}^N(t) = M_{k,i}^N(t) + R_{k,i}^N(t), \quad (6.36)$$

where

$$R_{k,i}^N(t) = Z_{k,i}^N(t) - \bar{Z}_{k,i}^N(t).$$

We have already shown in (3.14) that

$$\mathbb{E}(\sup_{s \leq t} |R_{k,i}^N(s)|) \leq CtN^{-1/2}.$$

In (6.36), we have that $[M_{k,i}^N, M_{l,j}^N]_t = 0$ for $(k, i) \neq (l, j)$, since the martingales almost surely never jump at the same time. Moreover, $[M_{k,i}^N, M_{k,i}^N]_t = \bar{Z}_{k,i}^N(t)$. Finally, $\mathbb{E}[(M_{k,i}^N(t))^2] = \mathbb{E}(\bar{Z}_{k,i}^N(t)) = m_t^k$.

We obtain

$$\begin{aligned} (m_t^k)^{-1} \mathbb{E}|U_{k,1}^N(t)| &\leq (m_t^k)^{-1} \mathbb{E}|M_{k,1}^N| + (m_t^k)^{-1} CtN^{-1/2} \\ &\leq \left(m_t^k\right)^{-1/2} + \frac{Ct}{m_t^k N^{1/2}}. \end{aligned} \quad (6.37)$$

Now we use that $t/N \rightarrow 0$ and $\liminf_{t \rightarrow \infty} m_t^k/t = \alpha_k > 0$ to deduce that

$$\limsup_{N, t \rightarrow \infty} (m_t^k)^{1/2} \mathbb{E} \left[|Z_{k,1}^N(t)/m_t^k - 1| \right] = \limsup_{N, t \rightarrow \infty} (m_t^k)^{1/2} \left[\frac{\mathbb{E}|U_{k,1}^N(t)|}{(m_t^k)} \right] \leq 1,$$

implying item 1 of the theorem.

The proof finishes in the lines of the proof of Theorem 10 of [13]. For fixed k and $i = 1, \dots, \ell_k \leq N_k$, we write

$$(m_t^k)^{1/2} (Z_{k,i}^N(t)/m_t^k - 1) = (m_t^k)^{-1/2} M_{k,i}^N(t) + (m_t^k)^{-1/2} R_{k,i}^N(t).$$

But $\mathbb{E}((m_t^k)^{-1/2} |\bar{R}_k^N(t)|) \leq CtN^{-1/2} (m_t^k)^{-1/2} \rightarrow 0$ as $t, N \rightarrow \infty$. So we only have to prove that $\left(((m_t^1)^{-1/2} M_{1,i}^N(t))_{1 \leq i \leq \ell_1}, \dots, ((m_t^n)^{-1/2} M_{n,i}^N(t))_{1 \leq i \leq \ell_n} \right)$ tends in law to $\mathcal{N}(0, I_{\ell_1 + \dots + \ell_n})$ which follows as in [13], proof of Theorem 10.

•

Proof of Theorem 2, supercritical case. We use the same notation as in the proof of the subcritical case and obtain in a first step the following control for (6.37).

$$(m_t^k)^{-1} \mathbb{E}|U_{k,1}^N(t)| \leq \left(m_t^k\right)^{-1/2} + \frac{Ce^{\alpha_0 t}}{m_t^k N^{1/2}}. \quad (6.38)$$

Therefore, for $N, t \rightarrow \infty$ under the constraint that $N^{-1/2}t^{-1}e^{\alpha_0 t} \rightarrow 0$,

$$\limsup_{N, t \rightarrow \infty} (m_t^k)^{1/2} \mathbb{E} \left[|Z_{k,1}^N(t)/m_t^k - 1| \right] = \limsup_{N, t \rightarrow \infty} (m_t^k)^{1/2} \left[\frac{\mathbb{E}|U_{k,1}^N(t)|}{(m_t^k)} \right] \leq C,$$

implying item 1 of the theorem.

The second item of the theorem follows using the same arguments as in the subcritical case. •

6.3 Proof of Theorem 4

The proof of Theorem 4 is based on the following steps. First, a standard calculus shows that we have an approximation result for the generators.

Lemma 2 *Grant the conditions of Theorem 4. Then there exists a constant C such that for all $\varphi \in C_b^3(\mathbb{R}^\kappa, \mathbb{R})$,*

$$\|A^X \varphi - A^Y \varphi\|_\infty \leq C \frac{\|\varphi\|_{3,\infty}}{N^2}.$$

The proof of the above lemma is straightforward and therefore omitted. In a next step, we obtain, applying Itô's formula with jumps twice, the following estimate.

Lemma 3 *Grant the conditions of Theorem 4. Then there exists a constant C such that for all $\varphi \in C_b^4(\mathbb{R}^\kappa, \mathbb{R})$, for any $\delta > 0$,*

$$\|P_\delta^Y \varphi - \varphi - \delta A^Y \varphi\|_\infty \leq C \delta^2 \|\varphi\|_{4,\infty}$$

and

$$\|P_\delta^X \varphi - \varphi - \delta A^X \varphi\|_\infty \leq C \delta^2 \|\varphi\|_{2,\infty}.$$

Again, the proof of the above lemma is straightforward and therefore omitted. Finally, we will use the following fact.

Lemma 4 *Grant the conditions of Theorem 4. Then there exists a constant C such that for all $\varphi \in C_b^4(\mathbb{R}^\kappa, \mathbb{R})$, for any $t > 0$,*

$$\|P_t^Y \varphi\|_{4,\infty} \leq C \|\varphi\|_{4,\infty}.$$

Proof By Kunita (1990) [29], see also Ikeda-Watanabe (1989) [25], there exists a version $\Phi_t(x)$ of the stochastic flow associated to the SDE (5.26) such that $Y^N(t) = \Phi_t(x)$, where $Y^N(t)$ is the solution of (5.26) starting from $Y^N(0) = x$. Under our assumptions, this flow is a flow of C^4 -diffeomorphisms. Then we can write $P_t^Y \varphi(x) = \mathbb{E}[\varphi(\Phi_t(x))]$, and by dominated convergence, for any $1 \leq k \leq \kappa$, $\partial_{x^k} P_t^Y \varphi(x) = \mathbb{E}[\sum_l \partial_{x^l} \varphi(\Phi_t(x)) \frac{\partial \Phi_t^l}{\partial x^k}(x)]$. The assertion follows then by iterating this argument, using classical estimates on the derivatives $\partial^\alpha \Phi_t(x)$ obtained e.g. in [25]. •

We are now able to finish the proof of Theorem 4.

Proof of Theorem 4 Fix $\delta > 0$ and write $t_k = k\delta \wedge t$, for $k \geq 0$. A standard trick yields

$$\|P_t^X \varphi - P_t^Y \varphi\|_\infty \leq \sum_{k:k\delta \leq t} \|P_{t-t_k+1}^X \Delta_\delta P_{t_k}^Y \varphi\|_\infty \leq \sum_{k:k\delta \leq t} \|\Delta_\delta P_{t_k}^Y \varphi\|_\infty, \quad (6.39)$$

where we define $\Delta_\delta \varphi(x) = P_\delta^X \varphi(x) - P_\delta^Y \varphi(x)$. Since $\varphi \in C_b^4$, also $P_{t_k}^Y \varphi \in C_b^4$.

By Lemma 3,

$$\|P_\delta^X \varphi(x) - P_\delta^Y \varphi(x)\|_\infty \leq \delta \|A^X \varphi - A^Y \varphi\|_\infty + \delta^2 \|\varphi\|_{4,\infty}.$$

Using Lemma 2, we deduce that

$$\|P_\delta^X \varphi(x) - P_\delta^Y \varphi(x)\|_\infty \leq [\delta C \frac{1}{N^2} + \delta^2] \|\varphi\|_{4,\infty}.$$

Together with Lemma 4 and (6.39), this yields

$$\|P_t^X \varphi - P_t^Y \varphi\|_\infty \leq C(\frac{1}{N^2} + \delta) \|\varphi\|_{4,\infty} \left(\sum_{k:k\delta \leq t} \delta \right) \leq Ct \frac{1}{N^2} \|\varphi\|_{4,\infty},$$

where the last inequality follows by choosing $\delta = \frac{1}{N^2}$ and using that $\text{card}\{k : k\delta \leq t\} \leq \frac{t}{\delta}$. •

6.4 Control theorem and proof of Proposition 6

We will use the control theorem which goes back to Strook and Varadhan (1972) [37], see also Millet and Sanz-Sole (1994) [33], theorem 3.5, in order to prove Proposition 6.

For some time horizon $T_1 < \infty$ which is arbitrary but fixed, write \mathbf{H} for the Cameron-Martin space of measurable functions $\mathbf{h} : [0, T_1] \rightarrow \mathbb{R}^2$ having absolutely continuous components $\mathbf{h}^\ell(t) = \int_0^t \dot{\mathbf{h}}^\ell(s) ds$ with $\int_0^{T_1} [\dot{\mathbf{h}}^\ell]^2(s) ds < \infty$, $1 \leq \ell \leq 2$. For $x \in \mathbb{R}^\kappa$ and $\mathbf{h} \in \mathbf{H}$, consider the deterministic system

$$\varphi = \varphi^{(N, \mathbf{h}, x)} \text{ solution to } d\varphi(t) = b(\varphi(t))dt + \frac{1}{\sqrt{N}} \sigma(\varphi(t)) \dot{\mathbf{h}}(t) dt, \text{ with } \varphi(0) = x, \quad (6.40)$$

on $[0, T_1]$. Thus φ is a function $[0, T_1] \rightarrow \mathbb{R}^\kappa$.

Using localization techniques as in Höpfner, Löcherbach and Thieullen (2015) [24], Theorem 5, we obtain the following result.

Proposition 8 *Grant the assumptions of Theorem 3. Denote by $Q_x^{t_1}$ the law of the solution $(Y^N(t))_{0 \leq t \leq t_1}$ of (5.26), starting from $Y^N(0) = x$. Let $\varphi = \varphi^{(N, \mathbf{h}, x)}$ denote a solution to*

$$d\varphi(t) = b(\varphi(t)) dt + \frac{1}{\sqrt{N}} \sigma(\varphi(t)) \dot{\mathbf{h}}(t) dt, \quad \varphi(0) = x.$$

Fix $x \in K$ and $\mathbf{h} \in \mathbf{H}$ such that $\varphi = \varphi^{(N, \mathbf{h}, x)}$ exists on some time interval $[0, T_1]$ for $T_1 > t_1$. Then

$$\left(\varphi^{(N, \mathbf{h}, x)} \right)_{|[0, t_1]} \in \overline{\text{supp} \left(Q_x^{t_1} \right)}.$$

We now show how to use Proposition 8 in order to prove Proposition 6.

Proof of Proposition 6 Fix $x \in \mathbb{R}^\kappa$ and $t_1 > 1$. Recall that, to simplify notation, we write $x = (x^1, \dots, x^\kappa)$, where $\kappa = 2 + \eta_1 + \eta_2$ instead of $x = (x^{1,0}, \dots, x^{1,\eta_1}, x^{2,0}, \dots, x^{2,\eta_2})$. In particular, the coordinates x^{η_1+1} and x^κ correspond to the two coordinates which are driven by Brownian noise. Let $\Gamma(t), t \in [0, T]$, be a parametrization of the periodic orbit, with T the periodicity of Γ .

Now, we choose a \mathcal{C}^∞ -function $\gamma = (\gamma^1, \gamma^2) : \mathbb{R}_+ \rightarrow \mathbb{R}^2$ satisfying

$$\begin{cases} \gamma^1(0) = x^1 \\ \gamma^2(0) = x^{\eta_1+2} \\ \gamma(s) \equiv (\Gamma^1(s), \Gamma^{\eta_1+2}(s)) \text{ on } [1, \infty). \end{cases} \quad (6.41)$$

We want to use γ as a smooth trajectory driving the two components φ^1 corresponding to $Y_{1,0}^N$ and φ^{η_1+2} corresponding to $Y_{2,0}^N$ from their initial position to a position on the periodic orbit, during a time period of length one.

We now show that it is indeed possible to choose a control \mathbf{h} such that $\varphi^1 = \gamma^1$ and $\varphi^{\eta_1+2} = \gamma^2$. Recall that the diffusion coefficient σ is null on every coordinate except the coordinates $\eta_1 + 1$ and κ . As a consequence, any choice of \mathbf{h} does only allow to influence directly these two coordinates $\eta_1 + 1$ and κ . However, above we have prescribed a trajectory γ to the two coordinates 1 and $\eta_1 + 2$. So we have to prove that such a choice of \mathbf{h} is possible.

Suppose for a moment that we have already found this control \mathbf{h} . Then, by the structure of b , once φ^1 and φ^{η_1+2} are fixed, we necessarily have

$$\varphi^2(t) = \frac{d\varphi^1(t)}{dt} + \nu_1 \varphi^1(t), \dots, \varphi^{\eta_1+1}(t) = \frac{d\varphi^{\eta_1}(t)}{dt} + \nu_1 \varphi^{\eta_1}(t) \quad (6.42)$$

for the first population, and

$$\varphi^{\eta_1+3}(t) = \frac{d\varphi^{\eta_1+2}(t)}{dt} + \nu_2 \varphi^{\eta_1+2}(t), \dots, \varphi^\kappa(t) = \frac{d\varphi^{\kappa-1}(t)}{dt} + \nu_2 \varphi^{\kappa-1}(t). \quad (6.43)$$

In other words, once φ^1 and φ^{η_1+2} are fixed, all other coordinates are entirely determined as measurable functions of γ . Moreover, by the structure of the equations given in (4.19), it is clear that for all $t \geq 1$, $\varphi(t) = \Gamma(t)$, i.e. the trajectory evolves on the orbit after time 1.

We have to show that we can indeed find a function \mathbf{h} which allows for the above choice of φ . The control \mathbf{h} is related to the coordinates φ^{η_1+1} and φ^κ through

$$\frac{d\varphi^{\eta_1+1}(t)}{dt} = -\nu_1 \varphi^{\eta_1+1}(t) + c_1 f_2(\varphi^{\eta_1+2}(t)) + \frac{c_1}{\sqrt{p_2}} \sqrt{f_2(\varphi^{\eta_1+2}(t))} \dot{\mathbf{h}}^1(t)$$

and

$$\frac{d\varphi^\kappa(t)}{dt} = -\nu_2 \varphi^\kappa(t) + c_2 f_1(\varphi^1(t)) + \frac{c_2}{\sqrt{p_1}} \sqrt{f_1(\varphi^1(t))} \dot{\mathbf{h}}^2(t).$$

In the above two formulas, all functions φ are known as measurable functions of the prescribed trajectory γ , and $\dot{\mathbf{h}}^1(t)$ and $\dot{\mathbf{h}}^2(t)$ have to be chosen. But since f_1 and f_2 are strictly positive, there exists indeed $\mathbf{h} : [0, \infty[\rightarrow \mathbb{R}^2$ achieving this choice of γ . It suffices to choose

$$\dot{\mathbf{h}}^1(t) = \frac{\frac{d\varphi^{\eta_1+1}(t)}{dt} + \nu_1 \varphi^{\eta_1+1}(t) - c_1 f_2(\varphi^{\eta_1+2}(t))}{c_1 / \sqrt{p_2} \sqrt{f_2(\varphi^{\eta_1+2}(t))}} \quad (6.44)$$

and

$$\dot{\mathbf{h}}^2(t) = \frac{\frac{d\varphi^\kappa(t)}{dt} + \nu_1\varphi^\kappa(t) - c_2f_1(\varphi^1(t))}{c_2/\sqrt{p_1}\sqrt{f_1(\varphi^1(t))}}. \quad (6.45)$$

Since f_1 and f_2 are strictly positive (and even lower bounded on K), $\dot{\mathbf{h}}(t)$ is well-defined. Now, notice that for all $t \geq 1$, $\varphi^1(t) = \Gamma^1(t)$ and $\varphi^{\eta_1+2}(t) = \Gamma^{\eta_1+2}(t)$ are evolving on the periodic orbit. Hence by (6.42) and (6.43), necessarily $\varphi(t) = \Gamma(t)$ for all $t \geq 1$. In particular, $\dot{\mathbf{h}}^1(t) = \dot{\mathbf{h}}^2(t) = 0$ for all $t > 1$.

Hence, we have constructed a control forcing the trajectory to be on the periodic orbit after a fixed time. By construction, $S(\varepsilon, \varphi) := \{\psi \in C(\mathbb{R}_+; \mathbb{R}^\kappa) : |\varphi(t) - \psi(t)| < \varepsilon \forall 1 \leq t \leq t_1\} \subset O$. Moreover, by Proposition 8, $\varphi|_{[0, t_1]} \in \text{supp}(Q_x^{t_1})$, whence $Q_x^{t_1}(S(\varepsilon, \varphi)) > 0$ and thus a fortiori $Q_x^{t_1}(O) > 0$, implying the assertion for all $t_1 > 1$.

•

References

- [1] ATHREYA, K.B., MURTHY, K.R. Feller's renewal theorem for systems of renewal equations. *J. of the Indian Institute of Science* 58, 10 (1976) 437–459.
- [2] BENAÏM, M., HIRSCH, M.W. Mixed Equilibria and Dynamical Systems arising from Fictitious Play in Perturbed Games. *Games and Econom. Behaviour*, 29 (1999) 36–72.
- [3] BERG, R.W., ALABURDA, A., HOUNSGAARD, J. Balanced Inhibition and Excitation Drive Spike Activity in Spinal Half-Centers. *Science*, 315 (2007) 390–393.
- [4] BERG, R.W., DITLEVSEN, S., HOUNSGAARD, J. Intense Synaptic Activity Enhances Temporal Resolution in Spinal Motoneurons *PLoS ONE*, 3 (2008) e3218.
- [5] BERG, R.W., DITLEVSEN, S. Synaptic inhibition and excitation estimated via the time constant of membrane potential fluctuations. *J Neurophysiol*, 110 (2013) 1021–1034.
- [6] BRÉMAUD, P., MASSOULIÉ, L. Stability of nonlinear Hawkes processes. *The Annals of Probability*, 24(3) (1996) 1563–1588.
- [7] CHEVALLIER, J. Mean-field limit of generalized Hawkes processes. Available on <http://arxiv.org/abs/1510.05620>, 2015.
- [8] CHEVALLIER, J., CACERES, M.J., DOUMIC, M., REYNAUD-BOURET, P. Microscopic approach of a time elapsed neural model. *Math. Mod. & Meth. Appl. Sci.*, 25(14) (2015) 2669–2719.
- [9] CHORNOBOY, E., SCHRAMM, L., KARR, A. Maximum likelihood identification of neural point process systems. *Biological Cybernetics* 59 (1988), 265–275.
- [10] COLLET, F., DAI PRA, P., FORMENTIN, M. Collective periodicity in mean-field models of cooperative behavior. *Nonlinear Differential Equations and Applications NoDEA*, 22(5) (2015) 1461–1482.

- [11] CRUMP, K. S. On systems of renewal equations. *J. Math. Analysis and Applications*, 30 (1970), 425–434.
- [12] DAI PRA, P., FISCHER, M., REGOLI, D. A Curie-Weiss Model with dissipation. *J. Stat. Phys.* 152 (2015), 37–53.
- [13] DELATTRE, S., FOURNIER, N., HOFFMANN, M. Hawkes processes on large networks. *Ann. App. Probab.* 26 (2016), 216–261.
- [14] DITLEVSEN, S., YIP, K.-P., HOLSTEIN-RATHLOU, N.-H. Parameter estimation in a stochastic model of the tubuloglomerular feedback mechanism in a rat nephron. *Math. Biosci.*, **194** (2005) 49–69.
- [15] DOUC, R., FORT, G., GUILLIN, A. Subgeometric rates of convergence of f -ergodic strong Markov processes. *Stochastic Processes Appl.*, 119(3) (2009) 897–923.
- [16] GALVES, A., LÖCHERBACH, E. Modeling networks of spiking neurons as interacting processes with memory of variable length. *Journal de la Société Française de Statistiques* 157 (2016), 17–32.
- [17] GLASS, L., MACKEY, M.C. From Clocks to Chaos: The Rhythms of Life. *Princeton University Press*, 1988.
- [18] GRAHAM, C. Chaoticity for multiclass systems and exchangeability within classes. *J. Appl. Prob.*, 45 (2008) 1196–1203.
- [19] GRAHAM, C., ROBERT, P. Interacting multi-class transmissions in large stochastic systems. *Ann. Appl. Prob.* 19, 6 (2009) 2334–2361.
- [20] HANSEN, N., REYNAUD-BOURET, P., RIVOIRARD, V. Lasso and probabilistic inequalities for multivariate point processes. *Bernoulli*, 21(1) (2015) 83–143.
- [21] HAWKES, A. G. Spectra of Some Self-Exciting and Mutually Exciting Point Processes. *Biometrika*, 58 (1971) 83–90.
- [22] HAWKES, A. G. AND OAKES, D. A cluster process representation of a self-exciting process. *J. Appl. Prob.*, 11 (1974) 93–503.
- [23] HIRSCH, M.W. Systems of Differential Equations which are Competitive or Cooperative, III: Competing Species. *Nonlinearity*, 1 (1988) 51–71.
- [24] HÖPFNER, R., LÖCHERBACH, E. AND THIEULLEN, M. Ergodicity and limit theorems for degenerate diffusions with time periodic drift. Application to a stochastic Hodgkin-Huxley model. Available on <http://arxiv.org/abs/1503.01648>, 2015.
- [25] IKEDA, N. AND WATANABE, S. Stochastic differential equations and diffusion processes. *North Holland* 1989.
- [26] ISHIHARA, K., KUNITA, H. A Classification of the Second Order Degenerate Elliptic Operators and its Probabilistic Characterization. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 30 (1974) 235–254.
- [27] JAHN, P., BERG, R.W., HOUNSGAARD, J., DITLEVSEN, S. Motoneuron membrane potentials follow a time inhomogeneous jump diffusion process. *J Comput Neurosci*, 31 (2011) 563–579.

- [28] KLOEDEN, P.E. AND LORENZ, J. Stable attracting sets in dynamical systems and in their one-step discretizations. *SIAM J. Numer. Anal.*, 23 (1986) 986–995.
- [29] KUNITA, H. Stochastic flows and stochastic differential equations. *Cambridge University Press* 1990.
- [30] MALLET-PARET, J., SMITH, H.L. The Poincaré-Bendixson Theorem for Monotone Cyclic Feedback Systems. *J. of Dynamics and Diff. Equations* 2, 4 (1990) 367–421.
- [31] SCHEUTZOW, M. Some examples of nonlinear diffusion processes having a time-periodic law. *The Annals of Probability*, 13(2) (1985) 379–384.
- [32] SCHEUTZOW, M. Noise can create periodic behavior and stabilize nonlinear diffusions. *Stochastic Processes Appl.*, 20 (1985) 323–331.
- [33] MILLET, A., SANZ-SOLÉ, M. A simple proof of the support theorem for diffusion processes. *Séminaire de Probabilités (Strasbourg)*, 28 (1994) 26–48.
- [34] REYNAUD-BOURET, P., RIVOIRARD, V., GRAMMONT, F., TULEAU-MALOT, C. Goodness-of-fit tests and nonparametric adaptive estimation for spike train analysis. *The Journal of Mathematical Neuroscience (JMN)* 4 (2014), 1–41.
- [35] SKELDON, A.C., PURVEY, I. The effect of different forms for the delay in a model of the nephron. *Math. Biosci. Eng.*, 2(1) (2005) 97–109.
- [36] SZNITMAN, A.-S. Topics in propagation of chaos. In *École d’Été de Probabilités de Saint-Flour XIX—1989*, vol. 1464 of *Lecture Notes in Math.* Springer, Berlin, 1991, 165–251.
- [37] STROOK, D., VARADHAN, S. On the support of diffusion processes with applications to the strong maximum principle. *Proc. 6th Berkeley Symp. Math. Stat. Prob. III*, pp. 333–359 (1972).
- [38] YOSHIKAWA, T. Stability theory by Ljapunov’s second method. *Publications of the Mathematical Society of Japan. Vol. 9. Tokyo. The Mathematical Society of Japan. VIII*, 1966.

Acknowledgments

This research has been conducted as part of the project Labex MME-DII (ANR11-LBX-0023-01). Author EL thanks Michel Benaïm for many stimulating discussions concerning monotone cyclic feedback systems. We thank Mads Bonde Raad for careful reading and critical remarks. We also thank two anonymous referees for helpful comments and suggestions. The work is part of the Dynamical Systems Interdisciplinary Network, University of Copenhagen. Moreover, it is part of the activities of FAPESP Research, Dissemination and Innovation Center for Neuromathematics (grant 2013/07699-0, S. Paulo Research Foundation).